

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ХАРКІВСЬКА НАЦІОНАЛЬНА АКАДЕМІЯ МІСЬКОГО
ГОСПОДАРСТВА

С. В. Сніжко

КОНСПЕКТ ЛЕКЦІЙ

з курсу

«СТАТИСТИКА»

*(для студентів 2 курсу денної та 3 курсу заочної форми навчання
освітньо-кваліфікаційного рівня бакалавр,
галузі знань 0305 «Економіка і підприємництво»,
напрямку підготовки 6.030509 «Облік і аудит»)*

Харків ХНАМГ 2010

Сніжко С. В. Конспект лекцій з курсу «Статистика» (для студентів 2 курсу денної та 3 курсу заочної форми навчання освітньо-кваліфікаційного рівня бакалавр, галузі знань 0305 «Економіка і підприємництво», напряму підготовки 6.030509 «Облік і аудит») / Авт.: С. В. Сніжко; Харк. нац. акад. міськ. госп-ва – Х.: ХНАМГ, 2010 – 196 с., укр. мов.

Автор: доц., к.е.н. С. В. Сніжко.

Рецензент: проф., д.е.н. Г. В. Ковалевський

Рекомендовано кафедрою менеджменту і маркетингу в міському господарстві, протокол № 04 від «28» жовтня 2009 р.

ЗМІСТ

ПЕРЕДМОВА	4
ЗМІСТОВИЙ МОДУЛЬ 1. ТЕОРЕТИЧНІ ОСНОВИ СТАТИСТИКИ.....	6
ТЕМА 1. ПРЕДМЕТ І МЕТОД СТАТИСТИКИ.....	6
ТЕМА 2. СТАТИСТИЧНЕ СПОСТЕРЕЖЕННЯ	21
ТЕМА 3. ЗВЕДЕННЯ І ГРУПУВАННЯ СТАТИСТИЧНИХ ДАНИХ.....	38
ТЕМА 4. СТАТИСТИЧНІ ПОКАЗНИКИ	50
ТЕМА 5. АНАЛІЗ РЯДІВ РОЗПОДІЛУ	69
ТЕМА 6. ВИБІРКОВИЙ МЕТОД.....	89
ЗМІСТОВИЙ МОДУЛЬ 1.2. МЕТОДИ СТАТИСТИКИ	106
ТЕМА 7. СТАТИСТИЧНА ПЕРЕВІРКА ГІПОТЕЗ	106
ТЕМА 8. СТАТИСТИЧНІ МЕТОДИ АНАЛІЗУ КОРЕЛЯЦІЙНИХ ЗВ'ЯЗКІВ.....	121
ТЕМА 9. АНАЛІЗ ІНТЕНСИВНОСТІ ДИНАМІКИ.....	139
ТЕМА 10. АНАЛІЗ ТЕНДЕНЦІЙ РОЗВИТКУ	152
ТЕМА 11. ІНДЕКСИ	164
ТЕМА 12. ПОДАННЯ РЕЗУЛЬТАТІВ СТАТИСТИЧНИХ ДОСЛІДЖЕНЬ	181
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ	193

ПЕРЕДМОВА

Статистична обізнаність є невід'ємною частиною професійної підготовки кожного майбутнього фахівця, що бачить себе в якості підприємця, фінансиста, економіста і, тим паче, керівника аудиторської фірми або головного бухгалтера великого підприємства – головного контролера-ревізора та «правої руки» першого керівника, розпорядника кредиту та «диспетчера» великих грошових потоків, помилка якого може призвести до фатальних для власників та багатотисячних колективів наслідків.

Сучасний стан фінансово-економічних відносин у суспільстві слід визнати вкрай недружелюбним по відношенню не тільки до державних інтересів, але й, в першу чергу, до малого та приватного бізнесу. Тому достеменний та чемний аналіз великої кількості формальної та спеціальної інформації, виявлення закономірностей та її систематизація дозволять уникнути похибок у процесі прийняття доленосних рішень.

Ця робота та, і це головне, її результати пов'язані із цільоспрямованим збиранням необхідної специфічної та спеціальної інформації, її зведенням і групуванням для зручного сприйняття. У подальшому отримана інформація повинна бути коректно проаналізована і узагальнена, на підставі чого повинні бути зроблені обґрунтовані висновки про існуючий стан виучуваного явища чи конкретної ситуації, напрямках їх розвитку, тенденціях, що склалися, наслідках ігнорування чи заходах щодо їх подолання. Крім того, фахівець повинен наочно та переконливо довести результати своїх досліджень їх замовнику або своєму керівникові.

Враховуючи наведене, суттєве значення має знайомство з основними категоріями, принципами, методологією та загальними визначеннями статистичного аналізу кількісної сторони масових соціально-економічних явищ, у нерозривному зв'язку з їх якісною стороною, конкретними умовами, місця і часу.

Основу статистичної грамотності в значній мірі надає вивчення запропонованого курсу «Статистика», що в більшій мірі відповідає загальноприйнятому курсу «Загальна теорія статистики». Основні постулати та методологія цієї науки та її застосування в господарчій практиці не змінюються протягом декількох століть, визначення якої перевірені практикою існування провідних економік світу, та які незалежні від соціально-економічних формацій, що склалися та конкурують на світовій арені.

Мова, якою розмовляють статистичні фахівці, – це «есперанто», яка зрозуміла для всіх зацікавлених і яка не залежить від юридичного поля, в якому вона застосовується. Багато в чому міжнародні стандарти бухгалтерського обліку побудовані на принципах єдиного розуміння статистичної методології – «фахівці розмовляють проміж собою мовою, яку розуміють усі причетні до справи».

Конспект лекцій «Статистика» розроблений на підставі:

- ГСВОУ МОНУ «Освітньо-кваліфікаційна характеристика бакалавра напряму підготовки 0501 – «Економіка і підприємництво», 04.06.2004 р. № 452 (з 2006 р. напрям 6.030509 «Облік і аудит»));
- СВО ХНАМГ Навчальний план підготовки бакалавра за спеціальністю 6.030509 «Облік і аудит», 2007 р.

Конспект лекцій «Статистика» ухвалений кафедрою «Менеджменту і маркетингу в міському господарстві» протокол № _____ від «___» _____ 2010 р.

ЗМІСТОВИЙ МОДУЛЬ 1. ТЕОРЕТИЧНІ ОСНОВИ СТАТИСТИКИ

ТЕМА 1. ПРЕДМЕТ І МЕТОД СТАТИСТИКИ

Питання для теоретичної підготовки

1. Поняття статистики та статистичного обліку. Класифікація дисциплін статистики.
2. Предмет статистичної науки. Історія й розвиток статистики.
3. Основні поняття в статистиці.
4. Загальна схема проведення статистичного дослідження.
5. Методи статистики. Статистична методологія.
6. Сучасна організація статистики в Україні.

1. Поняття статистики та статистичного обліку.

Класифікація дисциплін статистики

Термін «статистика» можна визнати одним з найбільш вживаних в економічній науці й практиці господарювання. У науковій літературі, спеціальних джерелах є маса визначень і тлумачень цього терміну. Узагальнюючи їх, слід зупинитися на наступному.

Спочатку він вживався у значенні «політичний стан». У науковий побут термін «статистика» був введений у XVIII ст. німецьким вченим Г. Ахенвалем (нім. *Statistik*, від італ. *stato*, лат. *status* – держава; інший переклад – положення, стан явищ, стан, положення речей).



Статистика – галузь знань (та відповідні їй навчальні дисципліни), у якій викладаються загальні питання збору, виміру й аналізу масових статистичних (кількісних або якісних) даних.

Велика радянська енциклопедія визначає **Статистику** як:

1. Вид суспільної діяльності, спрямований на одержання, обробку й аналіз інформації, що характеризує кількісні закономірності життя суспіль-

ства у всьому її різноманітті (техніко-економічні, соціально-економічні, соціально-політичні явища, культура) у нерозривному зв'язку з її якісним змістом.

У такому розумінні поняття **Статистики** збігається з поняттям статистичного обліку, що в умовах сучасного суспільства є провідним видом суспільного (народно-господарського) обліку. Визначальне значення статистики витікає з того, що уся інформація, яка має народно-господарську значимість і яка збирається шляхом бухгалтерського або оперативного обліку, в остаточному підсумку обробляється й аналізується за допомогою статистики. Вихідні методологічні принципи для побудови основних показників у всіх видах обліку є єдиними.

2. Галузь суспільних наук (та відповідні їй навчальні дисципліни), у якій викладаються загальні питання виміру й аналізу масових кількісних відносин і взаємозв'язків.



Статистичний облік – система реєстрації, узагальнення й вивчення масових, якісно однорідних соціально-економічних явищ у масштабі підприємства, галузі, економічного регіону або країни. Інформація статистичного обліку використовується органами влади й суб'єктами господарювання для прийняття управлінських рішень.

Узагальнюючи наведені визначення, варто підкреслити, що під статистикою розуміють три основні значення:

1) власне сукупності (масиви) показників, тобто набір певних статистичних даних (статистика народжуваності, статистика відвідувань сайту), що характеризують досліджувану сукупність явищ;

2) область практичної господарської діяльності зі створення й поширення особливої продукції – масових (статистичних) показників;

3) науку, що досліджує проблеми побудови, одержання й використання масових показників (розробку теорій показників, їхніх формул, методів аналізу тощо).

Статистика є багатогалузевою наукою й господарською практикою. Вона складається з окремих розділів, які, поєднуючи її самостійними частинами, тісно пов'язані між собою в єдине, нерозривне ціле.

На сьогодні немає загальноприйнятої класифікації розділів і відповідно дисциплін статистики. Однак, узагальнюючи досвід практичного застосування статистики, можна рекомендувати наступну класифікацію:

- *Загальна теорія статистики* (розглядає загальні основи теорії й практики статистики, включаючи використання її найважливіших методів).
- *Адміністративна статистика* (статистика управління державою, регіонами, містами, господарськими одиницями (підприємствами, організаціями, установами).
- *Економічна статистика*. Досліджує за допомогою статистичних показників масові явища й процеси, що відбуваються в економіці. Поєднують галузеві статистики економіки (статистику промисловості, будівництва, сільського господарства, транспорту, зв'язку, торгівлі та інше), а також статистики найважливіших сфер діяльності й показників економіки (фінансову статистику, біржову, статистику цін тощо).
- *Соціальна статистика* (статистика соціальної сфери діяльності, демографія, охорони здоров'я, освіта, культури, правова статистика тощо).
- *Статистика науки* (як галузі народного господарства).
- *Статистика природних ресурсів і навколишнього середовища*.
- *Міжнародна статистика*.
- *Математична статистика*.
- *Статистики природних наук*. Досліджують за допомогою статистичних показників масові явища й процеси, що відбуваються у природі (статистична фізика та ін.)
- *Інші області й дисципліни статистики*.

2. Предмет статистичної науки. Історія й розвиток статистики

Предмет і зміст статистичної науки довгий час були дискусійними. В результаті найбільш прийнятним вважається наступне:



Предметом статистики слід вважати кількісну сторону масових соціально-економічних явищ, в нерозривному зв'язку з їх якісною стороною, у конкретних умовах місця і часу.

З наведеного визначення впливають основні риси предмета статистичної науки:

- 1) Статистика – наука суспільна.
- 2) На відміну від інших суспільних наук статистика вивчає кількісну сторону суспільних явищ.
- 3) Статистика вивчає масові явища.
- 4) Статистика вивчає кількісну сторону явищ у нерозривному зв'язку з кількісною стороною і це знаходить своє втілення в існуванні системи статистичних показників.
- 5) Статистика вивчає кількісну сторону явищ у конкретних умовах місця й часу.

Професійна підготовка у менеджменті, економіці, бізнесі не буде повноцінною без вивчення професійної історії (історії професійних знань, умінь і практичних навиків). Нерідко історія того чи іншого питання стає ключем для вирішення найскладніших проблем.

Статистика – одна з найдавніших і найважливіших галузей людської діяльності. Ще в Древньому Єгипті, понад 5 тисяч років тому, проводилися статистичні переписи населення, вираховувалися індекси цін та інші статистичні показники. У роботах древньогрецького вченого Аристотеля (384 - 322 до н.е.) можна зустріти такі поняття, як «число», «система» й «фактор», які й у сучасній статистиці відіграють винятково важливу роль.

Перша статистична публікація – «Книга чисел» у Вітхому Завіті (Біблія). Найбільш ранні відомості про такі роботи в Китаї відносяться до 2238 р. до нашої ери. У Древньому Римі проводилися цензи (обліки) вільних громадян та їхнього майна; різноманітна практика обліково-статистичних робіт стала піддаватися теоретичним узагальненням.

Статистика як наука виникла у другій половині XVII століття.

Визнано, що основи статистичної науки закладені англійським економістом У. Петті (1623-1687) – винахідником політичної арифметики або статистики, автором першої у світі роботи, заснованої на економіко-статистичному методі аналізу («Політична арифметика», 1672 р.).

Досвід статистики знайшов відбиття у працях таких видатних вчених як німці М. Конрінг, Г. Ахенваль, А. Шлицер, росіяни М.В. Ломоносов, І.К. Кирилов, В.Н. Татіщев та інші.

У XVIII – XIX ст. у багатьох країнах світу виникли державні статистичні установи: Великобританія – 1700 р. (Парламентська статистика), Франція – 1789 р. (Статистичне бюро), Німеччина (Прусія) – 1805 р., Росія 1802 р. (Центральне статистичне бюро), США – 1866 р. (Статистичне бюро Міністерства торгівлі).

У XVIII – XX ст. сформувалися найважливіші методи математичної статистики – вибірковий метод і кореляційно-регресійний метод дослідження масових явищ і процесів.

Видатну роль у розвитку статистики в Україні зіграла Харківська економічна (статистична) школа, пов'язана з Харківським університетом (заснований у 1804 р.).

Розробка теорії динаміки циклічної ринкової економіки стала головним напрямком статистичних досліджень Харківської економічної школи за 200 років. У 1866 р. вперше у світі в дипломній роботі вихованця університету І.Г. Кауфмана «теорія коливання цін» була сформульована ідея про наявність економічних циклів у розвитку економіки. Дослідження у статистиці допомогли іншому вихованцеві Харківського університету

М. І. Туган-Барановському (1865 - 1919) створити теорію економічної кон'юнктури. М.І. Туган-Барановський був одним із засновників Української академії наук, Міністром фінансів в уряді Глави Центральної Ради України М.С. Грушевського. У сучасних умовах ринкової економіки статистика економічної і соціальної кон'юнктури відіграє значну роль у підвищенні економічної ефективності інституційних одиниць і якості життя населення.

У 1962 - 1964 рр. професор Харківського університету О.Г. Ліберман (1897 - 1982) пропонував розгорнути в СРСР ринкову економічну реформу. Ця реформа у формі « нової системи господарювання » (« хрущовська відлига ») слугувала прологом до « перебудови » М.С. Горбачова і до сучасних економічних реформ в Україні, Росії й інших країнах Східної і Центральної Європи.

Колишній студент Харківського університету С.А. Коваль (1901 - 1985) став єдиним українським лауреатом Нобелівської премії з економіки за статистичні дослідження економічного зростання держав залежно від обсягу та структури інвестицій.

Історія розвитку статистики свідчить, що статистична наука склалася в результаті теоретичного узагальнення накопиченого людством передового досвіду обліково-статистичних робіт. Викликано це, насамперед, потребами управління розвитком суспільства.

Таким чином, Статистика:

- це суспільна наука, що властивими тільки їй методами вивчає кількісну сторону масових явищ у нерозривному зв'язку з їхньою якісною стороною.
- дає числові вираження закономірностям розвитку суспільних явищ у конкретних умовах місця й часу.

3. Основні поняття в статистиці

Статистика як наука має свій предмет дослідження.



Предметом статистичного вивчення завжди виступають сукупності тих або інших множин однакісних, варіюючих явищ.

Статистична сукупність складається з одиниць сукупності. Кожна одиниця сукупності являє собою окремий випадок, прояву досліджуваної закономірності. Одиниці сукупності мають певні властивості, якості. Ці властивості прийнята називати ознаками.



Ознака (логічне) – це властивість, що відрізняє один об'єкт дослідження від усіх інших (*О. одинична*) або поєднує його в одну групу зі схожими з ним об'єктами (*О. загальна*).

Істотна ознака – властивість об'єкту дослідження, найбільш характерна для нього при розгляді його з відомої точки зору.

Дискретні ознаки – такі, що приймають тільки ізольовані значення, які відрізняються одна від одної на деяку скінченну величину (розряд робітників, кількість кімнат у квартирі).

Безперервні ознаки – значення, прийняті цими ознаками, можуть відрізнятися на яку завгодно малу величину (виробіток на одного робітника, рівень розрахунків за продукцію).

Іноді поняття ознаки досліджуваного явища ототожнюється з поняттям статистичний показник.



Статистичний показник – це узагальнююча знакова модель масового явища або процесу, що має кількісну (числову) сторону і якісну, змістовну сторону (найменування показника, його просторова й часова приналежність, одиниця виміру, алгоритм обчислення тощо).

Залежно від цілей статистичних показників їх можна підрозділити на два основних види:

- обліково-оціночні показники;
- аналітичні показники.



Обліково-оціночні показники – це статистична характеристика розміру якісно визначених соціально-економічних явищ в конкретних умовах місця й часу.

Аналітичні показники застосовуються для аналізу статистичної інформації і характеризують особливості розвитку досліджуваного явища (розрахункові показники, відносні й середні величини, показники динаміки й варіації, тісноти зв'язку тощо).

Треба мати на увазі, що в статистичних показниках виражається єдність якісної й кількісної сторін. Досліджувана ознака відображає лише якісну особливість досліджуваного явища.

Досліджувані статистичні ознаки можуть виражатися як якісними поняттями, так і числовими значеннями.

Ознаки розрізняють за способами їх виміру та іншими особливостями.

- Ознаки, виражені якісними поняттями прийнято називати **атрибутивними** або описовими (стать людини).
- Ознаки, виражені числовими значеннями називаються **кількісними** (вік людини)
- Ознаки, що приймають різні значення в окремих одиниць досліджуваного явища називаються **варіацією**.
- Значення ознаки, що варіює, у окремих одиниць досліджуваного явища називають **варіантом** (x_i).
- Частота зустрічаємості варіантів називається **частотою** (f_i).

- **Первинні ознаки** – характеризують одиницю сукупності в цілому. Це абсолютні величини.
- **Прямі ознаки** – це властивості, безпосередньо властиві тому об'єкту, що ними характеризується (чисельність робітників заводу).
- **Непрямі ознаки** – властивості, притаманні не самому об'єкту, а іншим сукупностям, що відносяться до об'єкту (робітники – оплата їх праці).
- Також виділяють **альтернативні ознаки**, які можуть приймати тільки одне з двох значень.
- **Дискретні ознаки** – це кількісні ознаки, що можуть приймати окремі значення, без проміжних значень між ними, частіше цілі числа (кількість кімнат в квартирі).
- **Безперервні ознаки** – точніше, беззупинно варіюючі ознаки – здатні приймати будь-які значення, звичайно, в окремих кордонах.
- **Моментні ознаки** характеризують досліджуваний об'єкт у якийсь момент часу, встановлений планом статистичного дослідження.
- **Інтервальні ознаки** – це ознаки, що характеризують результати процесів. Їх значення можуть виникати тільки за інтервал часу: рік, місяць, добу, але не на момент часу (число народжених за рік).

4. Загальна схема проведення статистичного дослідження

Загальною основою розробки і застосування статистичної методології є положення соціально-економічної теорії і принцип діалектичного методу пізнання явищ суспільного життя.

Відповідно до діалектичного методу пізнання статистика вивчає усі явища й процеси у їх взаємозв'язку і взаємозалежності, в русі й зміні, виділяючи їх різні типи і форми, встановлює нове і прогресивне та визначає напрямки розвитку. У своїх дослідженнях статистика спирається

на діалектичні категорії – випадкового і необхідного, одиночного і масового, індивідуального і загального.

Будь-яке статистичне дослідження систематизується за його цільовим призначенням і застосуванням та може бути зведене до чотирьох основних послідовних етапів (стадій) (рис.1.1):

- 1) статистичного спостереження;
- 2) зведення і групування матеріалів статистичного спостереження;
- 3) економічного аналізу та узагальнення даних, отриманих в результаті зведення і групування;
- 4) представлення результатів статистичного дослідження.

- **Перший етап:** *Статистичне спостереження* – планомірне, систематизоване збирання даних про явища суспільного життя шляхом реєстрації раніше визначених суттєвих ознак.

Мета – характеристика всього об'єкта дослідження.

Результат – статистичні дані про одиниці сукупності, які характеризують її з різних сторін.

- **Другий етап:** *Зведення і групування* – цифрове висвітлення явищ і процесів та групування їх по визначеним ознакам.

Результат – статистична сукупність, тобто множина об'єктів, явищ, східних за визначеними ознаками але маючих змінюючі характеристики (варіацію ознак).

- **Третій етап:** *Аналіз та узагальнення статистичної сукупності* за допомогою статистичних методів (середніх величин, кореляційно-регресійного аналізу, аналізу рядів динаміки, індексного та ін.).

Мета – визначення закономірностей в досліджуваних явищах, висновки про їх стан та тренди розвитку.

- **Четвертий етап:** *Представлення результатів статистичного дослідження* за допомогою статистичних таблиць і графіків.

Мета – наочність та доступність.

ЕТАПИ СТАТИСТИЧНИХ ДОСЛІДЖЕНЬ



Рис. 1.1 - Основні етапи статистичних досліджень

5. Методи статистики. Статистична методологія

Застосування у статистичних дослідженнях конкретних методів визначається поставленими при цьому завданнями і залежить від характеру вхідної інформації.

Метод (від грец. *μέθοδος* – «шлях крізь») – систематизована сукупність кроків, які необхідно зробити, щоб виконати певне завдання або досягти певної мети, спосіб збагнення істини.

Методологія (від грец. *μεθόδολογία* – вчення про метод) – система принципів і способів організації та побудови теоретичної й практичної діяльності людей; у той же час методологія є вченням про цю систему понять і їхніх відносин.

Методологію можна розглядати у двох зрізах: як теоретичну (і вона тісно пов'язана з розділом філософського знання гносеологією), так і практичну – орієнтовану на перетворення світу.

Для вивчення свого предмета статистика розробляє і застосовує різноманітні методи, сукупність яких утворює статистичну методологію.



Під **статистичною методологією** розуміється система принципів і методів їх реалізації спрямованих на вивчення кількісних закономірностей, що проявляються в структурі взаємозв'язків і динаміці соціально-економічних явищ.

Узагальнюючи досвід практичного застосування статистичних методів, можна рекомендувати наступну класифікацію:

1. Методи збору статистичної інформації (методи опитування, анкетування, звітних спостережень, самореєстрації, статистичних вибірок тощо).
2. Методи групування, сегментації, класифікацій і таблиць.
3. Методи середніх величин (методи побудови й використання середніх величин).

4. Графічні методи.
5. Балансові методи (включаючи баланси Системи національних рахунків, бюджети й баланси країн, регіонів, міст тощо).
6. Індексні методи й методи побудови рейтингів.
7. Матричні методи.
8. Методи кореляційно-регресійного аналізу.
9. Дисперсійний, факторний і компонентний аналіз.
10. Методи експертних оцінок.
11. Методи дослідження рядів динаміки.
12. Методи виробничих функцій.
13. Методи системного аналізу.
14. Методи фундаментального й технічного аналізу фінансової статистики.
15. Методи програмування (лінійного, нелінійного, оптимального, динамічного, дискретного, блокового тощо).
16. Методи пробних впроваджень (пробних статистичних переписів, цензів, кон'юнктурних обстежень тощо).
17. Методи контролю статистичної інформації (логічний і арифметичний контроль, повторні обстеження тощо).
18. Методи передачі й поширення статистичної інформації (у спеціальних статистичних публікаціях, періодиці, Інтернеті тощо).
19. Методи зберігання і захисту статистичної інформації (включаючи конфіденційну інформацію).
20. Організація державної, відомчої, регіональної, міської і біржової статистики України.

6. Сучасна організація статистики в Україні

Правовою основою державної статистичної діяльності є Конституція України. Закон України «Про державну статистику» регулює правові від-

носини в галузі державної статистики, визначає права і функції органів державної статистики, організаційні засади здійснення державної статистичної діяльності з метою отримання всебічної та об'єктивної статистичної інформації щодо економічної, соціальної, демографічної та екологічної ситуації в Україні та її регіонах і забезпечення нею держави та суспільства.



Державна статистика – централізована система збирання, опрацювання, аналізу, поширення, збереження, захисту та використання статистичної інформації.

Сьогодні головним обліково-статистичним центром в країні є Державний комітет статистики України (Держкомстат України) – спеціально уповноважений центральний орган виконавчої влади у галузі статистики. В його завдання входить подання офіційної статистичної інформації Президенту, Уряду, Парламенту, громадським і міжнародним організаціям, розробка науково обґрунтованої статистичної методології, координація статистичної діяльності регіональних органів виконавчої влади, аналіз економіко-статистичної інформації, складання національних рахунків і балансових розрахунків.

До системи органів державної статистики у відповідності з адміністративно-територіальним розподілом України входять територіальні органи державної статистики, що утворюються відповідно до законодавства спеціально уповноваженим центральним органом виконавчої влади у галузі статистики в Автономній Республіці Крим, областях, районах та містах.

Поряд із загальнодержавною статистикою існує відомча статистика, яка обслуговує підприємства, об'єднання, відомства і міністерства.

Враховуючи наведене, можна визначити **основні завдання статистики:**

- 1) систематичне спостереження і економічний аналіз матеріалів які характеризують хід виконання виробничих програм і подання їх

- у відповідні керівні структури для прийняття конкретних рішень;
- 2) вивчення і розробка цільових комплексних програм з соціально-економічних проблем конкретних регіонів і держави в цілому;
 - 3) удосконалення системи статистичних показників, які характеризують розвиток і результати соціально-економічних явищ і процесів;
 - 4) економічний аналіз стану і розвитку галузей економіки;
 - 5) вивчення чинників підвищення ефективності суспільного виробництва;
 - 6) створення загальнодержавної автоматизованої системи збору, обробки і передачі інформації;
 - 7) аналіз демографічних процесів.



Питання для самоконтролю по темі 1

1. Дайте визначення терміну «статистика».
2. Дайте визначення терміну «статистичний облік».
3. Наведіть класифікацію дисциплін статистики.
4. Кого можна вважати засновником статистичної науки?
5. Назвіть етапи розвитку харківської статистичної школи.
6. Дайте визначення терміну «предмет статистики».
7. Наведіть основні статистичні показники.
8. Наведіть основні етапи статистичного дослідження.
9. Дайте визначення терміну «статистична методологія».
10. Перелічте стадії статистичного дослідження.
11. Наведіть основні методи, які використовує статистика.
12. Наведіть визначення терміну «державна статистика».
13. Головні завдання державної і відомчої статистики.

ТЕМА 2. СТАТИСТИЧНЕ СПОСТЕРЕЖЕННЯ

Питання для теоретичної підготовки

1. Сутність статистичного спостереження і його завдання.
2. Основні організаційні форми статистичного спостереження. Види звітності.
3. Види й способи статистичного спостереження.
4. Методологічні й організаційні питання статистичного спостереження.
5. Помилки статистичного спостереження та види контролю його результатів.

1. Сутність статистичного спостереження і його завдання

Для отримання повних і вірогідних даних про стан будь-якого суспільного явища на певний момент часу, або про його розвиток за відповідний період часу проводять статистичне дослідження, яке складається з чотирьох послідовних етапів:

- 1) статистичного спостереження;
- 2) зведення і групування матеріалів статистичного спостереження;
- 3) економічного аналізу статистичної сукупності, отриманої в результаті зведення і групування;
- 4) наочного представлення результатів проведеного дослідження.

Етапи статистичного дослідження послідовні, пов'язані між собою, тому успіху можна досягти лише при добре підготовленій і організованій роботі на всіх його стадіях.



Статистичне спостереження – як перша стадія статистичного дослідження, являє собою планомірну, систематизовану, науково організовану роботу по збиранню і реєстрації масових первинних даних про явища і процеси суспільного життя у конкретних умовах місця і часу.

Статистичне спостереження являє собою науково-організовану роботу із збирання масових первинних даних про явища й процеси громадського життя. Будь-яке статистичне спостереження здійснюється за допомогою оцінки й реєстрації ознак одиниць досліджуваної сукупності у відповідних документах.



Ознака – відмітна риса, властивість, якість, властива одиниці сукупності, явища, процесу.

Ці дані в залежності від мети статистичного дослідження можуть бути різними за своїм змістом і способом отримання. Вони пізніше систематизуються, групуються, обробляються, аналізуються і узагальнюються.

Саме статистичне спостереження також складається з трьох етапів:

- а) підготовки спостереження;
- б) збирання матеріалу;
- в) контролю зібраного матеріалу.

На підготовчому етапі статистичного спостереження, відповідно до його мети і завдань, розробляються програма і організаційний план проведення спостереження. Вирішуються питання про зміст вихідної інформації, яким способом, якими засобами і в які терміни буде проведений облік фактів, як мусять бути організовані збирання і контроль отриманих первинних матеріалів. Повинні бути враховані також відповідні вимоги до оформлення цих матеріалів, яких вимагає техніка подальшої їх обробки на ПК із застосуванням спеціального програмного забезпечення.

Від якості статистичного спостереження залежить успіх всього статистичного дослідження. Статистичне спостереження повинно бути організоване таким чином, щоб в результаті його проведення були отримані об'єктивні, вірогідні, повні дані про досліджуване явище і, по можливості, в короткий термін. Це дасть змогу зробити правильні узагальнення і висновки.

Соціально-економічні явища і процеси, які спостерігаються, повинні мати наукову і практичну цінність та виражати їх типи.

Одним з важливих завдань статистичного спостереження є ретельна і всебічна перевірка якості зібраних матеріалів для забезпечення їх вірогідності.

Наукова організація статистичного спостереження потрібна для створення найкращих умов для отримання об'єктивно правильних матеріалів, які б давали змогу передбачити майбутні ситуації і робити обґрунтовані прогнози.

Статистичне спостереження проводять за строго визначеним **планом**, який включає програмно-методологічні і організаційні питання.

До *програмно-методологічних* відносять питання, зв'язані з розробкою програми спостереження, вивченням мети, об'єкта і одиниці спостереження, проектування формулярів і текстів інструкцій, встановленням джерел і способів збирання інформації.

До *організаційних* відносять питання про органи спостереження, терміни і місце проведення спостереження, складання попередніх списків одиниць досліджуваної сукупності, розставлення і підготовка кадрів та деякі інші.

Програма статистичного спостереження визначається правильно встановленими і конкретно сформульованими завданнями дослідження. Тому перш за все потрібно чітко сформулювати мету всієї роботи, а потім вирішувати усі інші питання програми спостереження.



Мета спостереження являє собою основний результат статистичного дослідження. Чітке і конкретне формулювання мети спостереження потрібне для того, щоб не збирати зайвих, непотрібних і неповних даних.

Завдання статистичного дослідження необхідно відобразити в статистичних показниках, для чого розробляють і складають макети кінцевих статистичних таблиць, в які заносять результати всієї роботи.



Об'єктом статистичного спостереження називається сукупність одиниць досліджуваного явища, про які повинні бути зібрані потрібні статистичні дані.

Визначивши об'єкт статистичного спостереження, потрібно вказати на його важливі ознаки і основні розпізнавальні риси, тобто встановити межі досліджуваної сукупності.

При періодичному обстеженні потрібно слідкувати, щоб досліджувана сукупність була більш менш однородною. Для цього статистика використовує **ценз** – обмежувальну ознаку, якій повинні задовольняти усі одиниці досліджуваної сукупності.

Поряд з визначенням об'єкта статистичного спостереження визначають також одиницю сукупності і одиницю спостереження.



Одиницею спостереження називають той первинний складовий елемент об'єкта статистичного спостереження, який є носієм ознак, що підлягають реєстрації.



Одиницею сукупності називається та первинна ланка, від якої отримують необхідні статистичні відомості.

Після того як визначені об'єкт, одиниця спостереження і одиниця сукупності, потрібно розробити зміст програми спостереження, що є основним питанням статистичного спостереження.



Програмою статистичного спостереження називається перелік чітко сформульованих питань, на які намічають отримати відповіді в процесі обстеження. Від якості її розробки залежать якість і цінність зібраного статистичного матеріалу.



Статистичним формуляром називається документ особливої форми куди збирають і записують відповіді на запитання програми спостереження.

Обов'язковим елементом статистичного формуляру є титульна і адресна частини, які необхідні для перевірки зібраних даних і їх наступного розроблення.

В титульній частині представляють назву статистичного спостереження, назву органу, який проводить спостереження, ким і коли затверджений формуляр, присвоєний йому номер.

В адресній частині записують точний адрес одиниці або сукупності одиниць спостереження і деякі інші відомості.

В практиці застосовують два види або дві системи статистичних формулярів – індивідуальну (карткову) і спискову.

Індивідуальним називається такий статистичний формуляр, в який заносять відомості про одну одиницю спостереження (листок обліку кадрів, одне підприємство, один робітник тощо).

Списковим називається такий статистичний формуляр, в якому реєструються відомості по декількох одиницях спостереження (відомість на заробітну плату, екзаменаційна відомість тощо).

Відповіді, які заносять до формулярів, виражаються словами, цифрами або у формі альтернативних відповідей (так чи ні).

Інструкцією називають сукупність роз'яснень і вказівок до програми статистичного спостереження. Вона повинна бути написана коротко і просто, вказівки повинні бути конкретними і чіткими.

2. Основні організаційні форми статистичного спостереження.

Види звітності

Уся різноманітність форм статистичного спостереження здійснюється у трьох основних формах:

- 1) у формі звітності підприємств, організацій і установ;
- 2) у формі спеціально організованого спостереження (перепис населення, інвентаризація майна, переоцінка основних фондів тощо);
- 3) у формі збору статистичної інформації з ділових документів.



Звітністю називають таку організацію статистичного спостереження, за якою відомості поступають в статистичні органи від підприємств, організацій і установ у вигляді обов'язкових звітів про їх діяльність в точно встановлені терміни.

Статистичну звітність складають на підставі даних первинного обліку.



Первинним обліком в статистиці називається ведення систематичних записів у формах первинних облікових документів про різні явища і процеси, які стосуються діяльності підприємств, організацій чи установ.

Звітність подають вищим організаціям і органам Державного комітету статистики в порядку, встановленому Державним комітетом статистики України щодо кожної форми.

У нашій країні розрізняють дві основні форми звітності:

- а) загальнодержавну;
- б) внутрівідомчу.



Загальнодержавна звітність є обов'язковою для усіх підприємств, організацій і установ. Вона подається уряду у зведеному вигляді міністерствами і відомствами безпосередньо або через Державний комітет статистики України.



Внутрівідомча звітність – це звітність, розроблена міністерствами і відомствами для своїх оперативно-господарських потреб.

Звітність у даний час є одним з основних джерел статистичної інформації про соціально-економічний розвиток держави.

За періодичністю звітність підрозділяють на:

- **поточну**, що складається і подається протягом року;
- **річну**, яка є найбільш повною за складом наведених показників.

Будь-яка форма звітності повинна містити реквізити (відомості):

- номер форми й дату затвердження;
- назва форми;
- назва підприємства, об'єднання, міністерства;
- коди класифікаторів;
- адреса або місцезнаходження даного підприємства;
- термін подання звіту;
- адреса, куди представляється звіт;
- підписи осіб, відповідальних за складання звітності.

Поряд із звітністю у практиці різних статистичних досліджень широко використовують спеціально організовані статистичні спостереження.



Спеціально-організованим статистичним спостереженням називається таке спостереження, яке проводиться із спеціальною метою на якусь дату для отримання інформації, котру, в силу певних причин, не можна зібрати із звітів, або для перевірки і уточнення даних звітності.

Одним з основних видів спеціально організованого спостереження є переписи.



Перепис – це спеціально організоване статистичне спостереження великого масштабу, яке охоплює усю країну або значну її частину, і проводиться одночасно за єдиною програмою. Його метою є визначення чисельності, складу, стану і розміщення досліджуваного об'єкта на встановлений критичний момент.

3. Види й способи статистичного спостереження

За повнотою охоплення спостереженням досліджуваного об'єкта розрізняють два його види – суцільне та несуцільне.

Суцільним називається таке спостереження, при якому обстеженню і реєстрації підлягають всі без винятку одиниці досліджуваної сукупності.

Несуцільним називається таке спостереження, при якому обстеженню і реєстрації підлягають не всі одиниці досліджуваної сукупності, а лише певна їх частина.

Несуцільні спостереження мають ту перевагу перед суцільним, що вони вимагають значно менших витрат сил і засобів, дозволяють застосувати докладнішу програму і досконаліший спосіб обліку фактів, швидше підводити підсумки обстеження і, отже, підвищують оперативність статистичних матеріалів.

Несуцільні спостереження в статистиці суттєво доповнюють основні матеріали, отримані в результаті суцільних спостережень. У багатьох випадках несуцільне спостереження є єдино можливим способом дослідження статистичної сукупності.

У практиці статистичної роботи застосовують наступні види несуцільного спостереження:

- 1) вибіркове спостереження;
- 2) монографічне спостереження;
- 3) метод основного масиву;
- 4) анкетне.

Вибірковим називається таке спостереження, при якому вся сукупність фактів характеризується за деякою її частиною, відібраною випадково. В його основі лежить випадковий відбір одиниць для обстеження, що гарантує незалежність результатів вибірки від волі осіб, які її проводять, і не допускає тенденційних помилок.

Монографічне спостереження являє собою детальне вивчення і опис окремого об'єкта або невеликої їх кількості за розширеною програмою. Таке спостереження проводиться з метою виявлення певних тенденцій і закономірностей розвитку явища або для вивчення і розпов-

судження передового досвіду окремих підприємств, організацій і установ. Воно також використовується для виявлення недоліків в роботі окремих підприємств з метою їх усунення і недопущення в майбутньому.

Метод основного масиву полягає у тому, що з усієї сукупності одиниць спостереженню підлягає переважна їх частка, в яку, як правило, попадають найбільш суттєві і крупні одиниці досліджуваної сукупності. Взяті разом вони мають значну питому вагу в сукупності за однією чи декількома основними для даного дослідження ознаками.

Анкетне спостереження ґрунтується на принципі добровільного заповнення адресатами надісланих або розданих їм спеціальних анкет з метою отримання потрібної для дослідження інформації. Недоліком анкетного спостереження є те, що перевірити достовірність зібраного матеріалу досить складно або неможливо. Його застосовують у випадках, коли не вимагається висока точність інформації, а лише наближені її характеристики.

За часом проведення статистичні спостереження поділяють на:

- а) поточне;
- б) періодичне;
- в) одноразове.

Поточним називається таке спостереження, яке ведеться систематично при безперервній реєстрації фактів в міру їх виникнення. Наприклад, реєстрація громадських актів (народження, смерть, шлюб, розлучення), облік виходів працівників на роботу, облік виробленої продукції на підприємстві тощо.

Періодичним називається таке спостереження, яке повторюється через певні, заздалегідь установлені рівні проміжки часу. Такі спостереження, як правило, характеризують стан явища на певний момент часу. Наприклад, щорічний перепис устаткування станом на 1 січня, облік чисельності працівників, товарних запасів, залишків матеріальних цінностей на 1 число кожного місяця тощо.

Одноразовим називається таке спостереження, яке проводиться в міру потреби один раз або час від часу, без дотримання точної періодичності (переписи виробничого устаткування, переписи автотракторної техніки тощо).

За способом збирання статистичних даних розрізняють:

- а) безпосереднє спостереження;
- б) документальне спостереження;
- в) опитування.

Безпосереднім називається таке спостереження, при якому самі реєстратори збирають потрібні дані шляхом особистих замірювань, зважувань і підрахунків одиниць об'єкта і на цій основі проводять записи у формулярі спостереження.

Документальним називається таке спостереження, при якому потрібні дані збирають і записують у формуляри на підставі використання різної документації.

Опитування – це таке спостереження, при якому відповіді на запитання записують зі слів опитуваної особи. Так проводять перепис населення.

У статистичній практиці використовують наступні *три способи опитування*:

- а) усне;
- б) самореєстрація;
- в) кореспондентський спосіб.

При *усному опитуванні* спеціально виділений працівник (реєстратор) розмовляє з опитуваною особою і з її слів сам заповнює формуляр.

При *самореєстрації* опитуваній особі вручають бланк обстеження, пояснюють питання, і опитувана особа сама заповнює формуляр. У визначений час обліковець збирає заповненні формуляри і перевіряє повноту і правильність їх заповнення.

Кореспондентський спосіб полягає в тому, що інформацію в органи, які проводять спостереження, надсилають добровільні кореспонденти, які

попередньо отримують від статистичних органів формуляри і інструкції щодо їх заповнення.

4. Методологічні й організаційні питання статистичного спостереження

З метою успішного проведення спостереження складають організаційний план його проведення.



Організаційний план – це основний документ, в якому зосереджено вирішення важливих питань організації і проведення статистичного спостереження.

Організаційний план статистичного спостереження містить дві групи питань:

- 1) програмно-методологічні;
- 2) організаційні.

До першої групи відносять питання, пов'язані з визначенням мети, об'єкта й одиниці спостереження, розробкою програми спостереження, проектуванням формулярів і тексту інструкцій, установленням джерел і способів збору даних.

Друга група включає питання про орган спостереження, термінах і місці проведення спостереження, складанні попередніх списків одиниць досліджуваної статистичної сукупності, розміщенню й підготовці кадрів тощо.

У програму спостереження повинні включатися тільки ті питання, які відповідають завданням дослідження, і на які можуть бути отримані правдиві, достовірні відповіді.

При організації статистичного спостереження повинні бути вирішені питання про час проведення спостереження, включаючи вибір сезону спостереження, встановлення терміну (періоду) і критичного моменту спостереження.



Під **періодом (терміном)** проведення спостереження розуміють час початку й закінчення збору відомостей.

Час спостереження – це час, до якого ставляться дані зібраної інформації.

Критичною називають дату, станом на яку повідомляються відомості.

До **організаційних питань** статистичного спостереження належать:

- визначення об'єкта, місця, часу і термінів спостереження;
- постановка мети і завдань спостереження;
- визначення органів спостереження;
- визначення прав і обов'язків окремих установ і організацій, які беруть участь у спостереженні;
- підготовчі роботи проведення спостереження;
- добір, навчання і інструктаж масових кадрів, потрібних для проведення спостереження;
- розмноження і розсилки формулярів спостереження;
- порядок здачі й приймання матеріалів спостереження;
- порядок отримання і подання попередніх і остаточних підсумків спостереження та інші практичні питання.

Організаційні плани складаються статистичними органами держави, починаючи з вищих і закінчуючи нижчими ланками.

Вищі статистичні органи головну увагу приділяють розв'язанню загальних організаційно-методологічних питань, таких як визначення завдань спостереження, його об'єкта, одиниці і термінів проведення тощо.

Нижчі ланки статистичних органів розв'язують головним чином конкретні організаційні завдання на місцях.

Статистичне спостереження у загальнодержавному масштабі організовує Державний комітет статистики України та його місцеві органи.

Міністерства, відомства, наукові та інші установи проводять статистичні спостереження в основному локального характеру.



Місце спостереження – це місце, де проводиться реєстрація фактів спостереження, які записуються у спеціальних формулярах.

Вибір місця спостереження повинен забезпечити повноту охоплення об'єкта спостереження, високу якість фіксації даних і простоту проведення спостереження.

5. Помилки статистичного спостереження та види контролю його результатів

Точність статистичного спостереження є важливою і основною вимогою органів державної статистики. Однак, хоч як би старанно не було підготовлене статистичне спостереження, в процесі його проведення трапляються помилки, які призводять до зниження його точності.



Точністю статистичного спостереження називають ступінь відповідності значення будь-якої ознаки, визначеної за допомогою статистичного спостереження, її дійсному значенню. Чим ближчі значення ознак, отриманих в результаті статистичного спостереження, до їх фактичних значень, тим точніше проведене спостереження.

Точність статистичного спостереження визначається як відношення даних спостереження до дійсних значень досліджуваних величин, або як різниця між ними.



Помилками спостереження називаються розходження між встановленими статистичним спостереженням і дійсними значеннями досліджуваних величин.

Помилки спостереження виникають внаслідок неточностей при збиранні і реєстрації значень досліджуваних ознак.

Недопущення і попередження помилок є одним з важливих завдань організації і проведення статистичного спостереження. Невірні статистичні дані можуть призвести до прорахунків в державному управлінні економікою, серйозних помилок в науковому плануванні і прогнозуванні та інших негативних наслідків. Тому в Україні законодавчо встановлено сувору відповідальність посадових осіб за навмисні викривлення статистичних даних.

Залежно від характеру, ступеню впливу на кінцеві результати, джерел і причин виникнення неточностей розрізняють наступні **типи помилок статистичного спостереження**:

- а) помилки реєстрації;
- б) помилки репрезентативності.

Кожний з цих типів помилок, у свою чергу, поділяється на випадкові (ненавмисні) й систематичні (навмисні).

Помилки реєстрації виникають внаслідок неправильного встановлення фактів у процесі спостереження, помилкового запису їх значень, або обох причин разом.

Випадковими називаються помилки реєстрації, які можуть виникати внаслідок різних випадкових причин. Наприклад, опитувана особа може обмовитись, а реєстратор недочути чи випадково переставити місцями цифри. Такі неточності діють у протилежних напрямках і при достатньо великій кількості спостережень взаємно погашаються.

Систематичні помилки реєстрації виникають внаслідок певних причин, діють в одному і тому ж напрямку і спричиняють серйозні викривлення загальних результатів статистичного спостереження.

Статистичні помилки реєстрації можуть бути наслідком свідомого викривлення фактів – це навмисні приписки або приховування у звітах фактичних даних.

Помилки реєстрації виникають як при суцільному, так і при несуцільному спостереженні.

На відміну від помилок реєстрації помилки репрезентативності властиві тільки несуцільному спостереженню.

Помилками репрезентативності називаються відхилення значень ознак відібраної і обстеженої частини сукупності від значень ознак всієї досліджуваної сукупності.

Випадкові помилки репрезентативності виникають внаслідок того, що відібрана випадковим, неупередженим способом частина досліджуваної сукупності недостатньо повно відтворює всю сукупність в цілому.

Систематичні помилки репрезентативності виникають внаслідок порушення принципів неупередженого, випадкового відбору одиниць для обстеження.

З метою отримання в процесі статистичного спостереження високоякісних матеріалів статистичні органи здійснюють постійний контроль за ходом проведення спостереження, систематично перевіряють стан первинного обліку і звітності на підприємствах, організаціях і установах.

Після закінчення спостереження матеріали, зібрані в процесі його проведення, старанно перевіряються за повнотою охоплення об'єкта спостереження, якістю заповнення формулярів і інших документів.

Статистика використовує два **способи контролю матеріалів спостереження**:

- а) арифметичний (лічильний);
- б) логічний.

Арифметичний контроль полягає в лічильній перевірці підсумкових даних звітів або формулярів і погодженні тих показників, які взаємозв'язані між собою і можуть бути виведені одні з одних. Наприклад, в шаховій таблиці будь-якого значення підсумки рядків і колонок повинні співпадати, а якщо такого співпаданя не має, тоді шукають помилку в рядках чи колонках.

Логічний контроль полягає в співставленні взаємозв'язаних між собою відповідей на питання формуляра статистичного спостереження і виясненні їх логічної сумісності. Якщо виявляють логічно несумісні відповіді, шляхом подальшого співставлення з відповідями на інші запитання встановлюють, яка з відповідей записана невірно.

Основною умовою успішного проведення будь-якого статистичного дослідження на всіх його стадіях, в тому числі на стадії спостереження, є висока якість зібраного матеріалу.



Питання для самоконтролю по темі 2

1. *Поняття про статистичне спостереження.*
2. *Наведіть етапи статистичного дослідження.*
3. *Завдання статистичного спостереження.*
4. *Мета статистичного спостереження.*
5. *Об'єкт статистичного спостереження.*
6. *Одиниця статистичного спостереження.*
7. *Програма статистичного спостереження.*
8. *Критичний момент статистичного спостереження.*
9. *Період статистичного спостереження.*
10. *Організаційний план статистичного спостереження.*
11. *Основні організаційні форми статистичного спостереження.*
12. *Поняття про статистичну звітність.*
13. *Загальнодержавна і відомча звітність.*
14. *Спеціально організоване статистичне спостереження.*
15. *Поняття про переписи.*
16. *Суцільні й несуцільні статистичні спостереження.*
17. *Поняття про вибіркове спостереження.*
18. *Точність статистичного спостереження.*
19. *Помилки статистичного спостереження.*
20. *Помилки реєстрації.*
21. *Випадкові і систематичні помилки реєстрації.*
22. *Помилки репрезентативності.*
23. *Способи контролю результатів статистичного спостереження.*

ТЕМА 3. ЗВЕДЕННЯ І ГРУПУВАННЯ СТАТИСТИЧНИХ ДАНИХ

Питання для теоретичної підготовки

1. Поняття про статистичне зведення.
2. Завдання і значення статистичних групувань, їхні види.
3. Основні правила і принципи утворення груп.
4. Види групувань.
5. Системи групувань за багатьма ознаками.
6. Необхідність створення системи групувань, та основні вимоги до них.

1. Поняття про статистичне зведення

Другим етапом *статистичного дослідження*, метою якого є надати зведену характеристику всієї сукупності фактів за допомогою узагальнюючих статистичних показників, слід вважати *зведення і групування статистичних матеріалів*.

Результатом проведення статистичного спостереження може бути велика кількість різноманітних відомостей про кожну одиницю досліджуваної сукупності. Для того, щоб на основі цих відомостей можна було зробити певні висновки, потрібно всю масу окремих даних привести до відповідного порядку, систематизувати й обробити.



Статистичним зведенням називається спеціальна обробка первинних даних статистичного спостереження з метою отримання узагальнюючих характеристик досліджуваного явища чи процесу за рядом суттєвих для них ознак.

Перед зведенням отриманого первинного матеріалу його потрібно логічно проконтролювати і прийняти. Попередній теоретичний аналіз необхідний для того, щоб під час зведення не губились основні риси досліджуваних явищ в загальних підсумках. Опрацьований матеріал перевіряється за повнотою охоплення обстежуваних одиниць і якістю отриманих про них даних.

Повнота і якість отриманої інформації перевіряється за допомогою логічного і лічильного контролю, виявлені дефекти виправляють. Важливою умовою своєчасного і правильного проведення статистичного зведення є суворе дотримання звітної дисципліни. Тільки після того, як весь первинний статистичний матеріал старанно проконтрольований і належним чином виправлений, можна приступати до його зведення.

Зведення може бути **просте** – як вузькотехнічна операція з підрахунку підсумків первинного статистичного матеріалу, а також **складне** – яке передбачає групування даних, розробку системи показників, підрахунок групових і загальних підсумків та виклад результатів зведення у вигляді статистичних таблиць чи графіків.

Статистичне зведення обов'язково проводять за наперед розробленою програмою, яка повинна відповідати завданням статистичного дослідження з врахуванням прийнятої форми організації зведення.

В залежності від завдань статистичного дослідження **програма зведення** встановлює групувальні ознаки, кількість груп та макети розроблювальних таблиць. Програма повинна бути складена таким чином, щоб в результаті зведення отримати матеріал, який характеризує досліджуване явище з різних його сторін.

Для успішного здійснення статистичного зведення складається *план* його проведення. План повинен містити розв'язання питань організації зведення: послідовність і терміни виконання окремих частин зведення, оформлення його результатів у вигляді таблиць, публікацій у вигляді статистичних збірників та ін.

За формою організації зведення може бути *централізованим* і *децентралізованим*.

При централізованій формі організації зведення всі матеріали спостереження обробляють і синтезують в Державному комітеті статистики України. Суттєвою перевагою даної форми зведення є те, що вона дає можливість його автоматизації і використання єдиної методології обробки

даних. При децентралізованій формі організації зведення матеріали спостереження обробляються і узагальнюються на місцях, а в центральні органи надсилають зведену інформацію по регіонах. Децентралізована форма зведення дещо дешевша і оперативніша за централізовану.

На практиці поєднують територіально-децентралізовану і централізовану форму зведення.

3.2. Завдання і значення статистичних групувань, їхні види

Під час обробки матеріалів статистичного дослідження виникає потреба *виділення якісно однорідних груп*, типів, а вже потім їх описання за допомогою відповідних кількісних характеристик. Застосування таких узагальнюючих показників як відносні і середні величини, індекси та інші, можливе лише після того, як статистичний матеріал розподілений на однорідні групи.



Метод групування у статистиці полягає в розчленуванні усіх одиниць досліджуваної сукупності на якісно однорідні групи за істотними для них ознаками. Використання групувань дозволяє вивчити зв'язки й залежність між явищами, що вивчаються, структуру й структурні зміни досліджуваної сукупності.

Групування є центральним моментом будь-якого зведення, завдяки чому матеріал статистичного спостереження приймає систематизований вигляд. Ознаки, покладені в основу групування, називаються *групувальними*. Групування одиниць досліджуваної сукупності за будь-якою ознакою дає можливість формування *рядів розподілу*.

Групувальні ознаки можуть мати кількісний вимір (наприклад, вік працівника, стаж роботи, заробітна плата та ін.), тому вони називаються кількісними, а ряди їх розподілу – **варіаційними рядами**.

Якщо групувальні ознаки відображають певні властивості одиниць

сукупності (наприклад, стать, національність, освіту тощо), вони називаються якісними, а ряди розподілу – **атрибутивними**.

При групуванні одиниць сукупності за територіальною ознакою отримують **географічні** або **територіальні** ряди розподілу. Вони дають уяву про розміщення або ступінь розповсюдження тих чи інших явищ в просторі.

Особливим видом групувань в статистиці є *класифікація*.



Класифікацією в статистиці називається стійке і фундаментальне групування одиниць сукупності за атрибутивною ознакою на подібні і відмінні групи і підгрупи. Перелік цих груп і підгруп розглядається як своєрідний статистичний стандарт, затверджений Державний комітетом статистики України.

Наприклад, класифікація галузей економіки, класифікація основних фондів, класифікація професій тощо. Статистичні класифікації ґрунтуються на таких суттєвих ознаках, які мало змінюються і існують тривалий час.

Групування, будучи першою сходинкою статистичного аналізу, одночасно є підготовчою стадією для більш глибокого аналізу досліджуваного статистичного матеріалу.

3.3. Основні правила і принципи утворення груп

Перед проведенням простих, а тим більше комбінованих групувань потрібно розв'язати питання про кількість груп, розмір інтервалів та ін.

При групуванні за *атрибутивними ознаками* число груп, на які ділять досліджувану сукупність, визначається кількістю різновидів цієї ознаки.

Окремим випадком атрибутивних групувань є **альтернативне групування**, коли є всього два варіанти атрибутивної ознаки, один з яких виключає інший. Наприклад, розподіл робітників підприємства на таких, що мають спеціальну освіту, і таких, які цієї освіти не мають, тощо.

Інший характер має **групування за кількісними ознаками**, при якому виникає питання про кількість груп, числові межі окремих груп та інтервали групування. Наприклад, групування робітників за стажом роботи, тарифним розрядом, чи заробітною платою; групування заводів за вартістю основних виробничих фондів, випуском або реалізацією продукції тощо.

При розв'язанні питання про те, скільки доцільно утворити груп, беруть до уваги варіацію ознаки і число спостережень. Чим інтенсивніше змінюється ознака і чим більша сукупність одиниць, тим більше число груп потрібно утворити.

Однак, як загальний принцип розв'язання питання про необхідну кількість груп, виступає вимога, щоб ця кількість була оптимальною і щоб до кожної групи потрапила достатньо велика кількість одиниць. При великій кількості груп відбудеться розпорошення одиниць досліджуваної сукупності, однорідні одиниці попадуть в різні групи. А при малій кількості груп до однієї й тієї ж групи попадуть одиниці різних типів.

При групуванні за кількісними ознаками виникає суттєве питання про *вибір розміру інтервалів групування*.



Інтервалом групування називається різниця між максимальним і мінімальним значеннями ознаки в кожній групі.

Інтервали в структурних і аналітичних групуваннях можуть бути рівними і нерівними залежно від характеру розподілу одиниць сукупності за даною ознакою. Нерівні інтервали, в свою чергу, можуть бути прогресивно-зростаючими або прогресивно-спадаючими.

Якщо варіація досліджуваної ознаки знаходиться у порівняно вузьких межах і розподіл близький до нормального, то застосовують **рівні інтервали**.

Величину інтервалу при групуванні з рівними інтервалами визначають шляхом ділення розмаху варіації на число груп за формулою:

$$i = \frac{x_{\max} - x_{\min}}{n}, \quad (3.1)$$

де x_{\max} – максимальна величина ознаки;

x_{\min} – мінімальна величина ознаки;

i – розмір інтервалу;

n – число груп.

Розмір інтервалу залежить від числа груп і варіації досліджуваної ознаки. Чим більшою буде варіація ознаки, тим більшим буде розмір інтервалу і чим більше число груп, тим менший розмір інтервалу.

Число груп орієнтовно можна визначити за формулою американського вченого Стерджеса:

$$n \approx 1 + 3,321 * \lg N, \quad (3.2)$$

де N – число одиниць сукупності, яка підлягає групуванню.

Тоді формула для визначення розміру інтервалу матиме вигляд:

$$i = \frac{x_{\max} - x_{\min}}{1 + 3,321 * \lg N}, \quad (3.3)$$

В соціально-економічній статистиці часто застосовують групування з **нерівними інтервалами**. Застосування нерівних інтервалів, які прогресивно збільшуються або зменшуються, зумовлено самою природою більшості соціально-економічних явищ, коли в нижчих групах велике значення мають навіть малі відмінності в показниках, а у вищих групах такі відмінності суттєвого значення не мають. Так, наприклад, для нижчих груп, при групуванні підприємств за чисельністю працюючих різниця в 50 чи 100 чоловік має велике значення, а для вищих груп, в яких зосереджені великі підприємства, така різниця не суттєва.

В статистичних групуваннях часто розмежовують дві якісно відмінні групи підприємств. Наприклад, підприємства, які не виконали план і такі, які виконали план на 100 % і більше.

Групування, метою яких є утворення якісно однорідних груп, використовують **спеціалізовані інтервали**. В таких групуваннях межа інтервалу встановлюється там, де відбувається перехід від однієї якості до іншої. Наприклад, групування дітей за віком, за характером відношення чоловічого населення до трудової діяльності тощо.

3.4. Види групувань

Серед багатьох завдань, які розв'язуються з допомогою статистичних групувань, можна виділити три основних:

- 1) розподіл всієї сукупності на якісно однорідні соціально-економічні типи;
- 2) вивчення структури явищ і структурних зрушень в них;
- 3) виявлення і характеристика взаємозв'язку між ознаками досліджуваного явища.

Основними видами групувань, за допомогою яких розв'язуються ці завдання, слід вважати наступні:

- **Типологічними** називаються групування, за допомогою яких проводять розподіл досліджуваного суспільного явища на класи або соціально-економічні типи. На основі всебічного теоретичного аналізу досліджуваної сукупності виділяють її головні й найхарактерніші типи або групи, вивчають істотні відмінності між ними, а також спільні ознаки для всіх груп.

За допомогою типологічних групувань вивчають розподіл підприємств за формами власності, поділ економіки на сферу матеріального виробництва і невиробничу сферу тощо.

Виділені в результаті типологічних групувань окремі типи явищ вивчаються за їх складом. Групування дозволяє також охарактеризувати структуру і структурні зрушення досліджуваного явища. Так, за допомогою групувань вивчають склад населення за віком, статтю, освітою, національністю, сімейними положеннями; склад працівників – за професіями, стажем роботи, віком та іншими ознаками.

- **Структурним** групуванням в статистиці називається розчленування якісно однорідної сукупності одиниць на групи, які характеризують її структуру і структурні зрушення за відповідний період часу.

Для вивчення процесу концентрації в промисловості підприємства групують за вартістю основних виробничих фондів, за середнь-списковою чисельністю працюючих, за обсягом виробленої чи реалізованої продукції, за рівнем механізації і автоматизації виробництва тощо. Кожна з наведених групувальних ознак по-своєму відображає концентрацію. Для трудомістких галузей економіки концентрацію відображають через групування за числом робітників, для енергомістких галузей – із допомогою групувань за енергозатратами, а в деяких випадках групування за однією ознакою доповнюють групуваннями за іншими ознаками.

Одним з основних завдань, яке розв'язується за допомогою статистичних групувань є дослідження взаємозв'язків варіаційних ознак в межах, як правило, однорідної сукупності.

- **Аналітичними** групуваннями в статистиці називаються такі, за допомогою яких виявляють і вивчають взаємозв'язок між окремими ознаками досліджуваного соціально-економічного явища.

Усі явища суспільного життя і їх ознаки пов'язані між собою і залежать одне від одного. Взаємозв'язані ознаки діляться на *факторні* й *результативні*.

Факторною називається ознака, під впливом якої змінюється залежна від неї інша ознака.

Результативною називається ознака, яка змінюється під впливом факторної ознаки.

Раніше наведені групування, зроблені на основі первинного статистичного матеріалу, називаються **первинними групуваннями**.

Але іноді, при статистичному дослідженні, виникає необхідність раніше згрупований матеріал перегрупувати. **Вторинним групуванням** в статистиці називається процес утворення нових груп на основі раніше проведеного групування первинних даних.

Метод вторинного групування використовується для утворення, на основі групування за кількісною ознакою, якісно однорідних груп або типів, приведення кількох групувань з різними інтервалами до одного виду, з метою порівняльності та утворення більш укрупнених груп, в яких чіткіше простежується характер розподілу. Його застосовують також при економічному аналізі роботи підприємств для приведення до порівняльного виду їх вже згрупованих даних, але не зіставлених за територією або періодами часу.

Статистика використовує два ***способи утворення нових груп***.

Перший, найбільш простий і розповсюджений, – це **спосіб зміни** (як правило укрупнення) інтервалів первинного групування. У більшості випадків тут виходять із передбачення про нормальний розподіл ознак всередині інтервалів. Цей спосіб вторинного групування застосовується для зведення двох і більше групувань з неоднаковими інтервалами до одного виду з метою зіставлення.

Другий спосіб вторинного групування базується на закріпленні за кожною групою повної частки одиниць сукупності (**спосіб дольового перегрупування**).

3.5. Системи групувань за багатьма ознаками

З метою більш сталого та комплексного дослідження процесів розвитку явищ суспільного життя в реальному часі виникає необхідність групування даних статистичних досліджень за двома і більше ознаками. Такі групування в статистиці називаються **складними**.

Найбільш розповсюдженим видом складних групувань слід вважати комбіновані групування, при яких розчленовані на групи сукупності піддаються подальшому дрібненню на підгрупи за іншими додатковими ознаками.

Комбінованими в статистиці називаються групування, коли утворені групи за однією ознакою, діляться потім на підгрупи за однією і більше ознаками, взятими в комбінації.

Одночасне використання декількох групувальних ознак дозволяє виявити і порівняти такі відомості і зв'язки між досліджуваними ознаками, які не можна знайти через ізольовані групування за рядом групувальних ознак. Як правило, на практиці при проведенні комбінованого групування обмежуються трьома - чотирма ознаками.

Для розв'язання цього завдання математичною статистикою розроблений **метод багатовимірних групувань – метод розпізнавання образів**. На основі розпізнавання образів віднесення до окремих груп проводять за комплексом ознак, які утворюють «ознаковий простір», де кожна ознака розглядається як координата.

Якщо в наборі сукупності є ряд одиниць з фіксованими ознаками, то кожна з цих одиниць розглядається як точка в багатовимірному просторі. Всі точки, розташовані в багатовимірному просторі, утворюють скупчення, які складають групи (таксони, кластери). Завданням виділення цих скупчень точок у просторі й займається теорія розпізнавання образів.

Геометрична близькість декількох точок в багатовимірному просторі означає подібність становища відповідних об'єктів, тобто їх кількісна однорідність за досліджуваними ознаками.

Одним з розповсюджених критеріїв міри близькості між об'єктами є евклідова віддаль:

$$d_{ij} = \sqrt{\sum_{k=1}^n (x_{ki} - x_{kj})^2}, \quad (3.4)$$

де x_{ki} – значення k -ої ознаки в i -му об'єкті;

x_{kj} – значення k -ої ознаки в j -му об'єкті.

При зменшенні цієї віддалі близькість зростає, а при її збільшенні – зменшується.

3.6. Необхідність створення системи групувань та основні вимоги до них

Завдання всебічного аналізу соціально-економічних явищ і процесів не розв'язується через складання якогось одного групування, яке б характеризувало типи, структуру і взаємозв'язки даного економічного явища. Такий аналіз вимагає складання системи групувань за багатьма ознаками, яка допоможе охарактеризувати розвиток певного явища в цілому. Утворені системи групувань повинні відповідати загальним методологічним вимогам і підпорядковуватись *логічним і формальним критеріям*.

Логічні критерії вимагають наступного:

- а) система повинна всебічно характеризувати досліджуваний об'єкт;
- б) системою групувань повинні розв'язуватись типологічні, структурні і аналітичні завдання дослідження;
- в) кожне окреме групування повинно бути однією з логічних ланок в загальній системі;
- г) висновки за кожним групуванням не повинні суперечити одні одним;
- д) система групувань повинна бути стабільною і не піддаватись частим змінам в часі.

Основними **формальними критеріями** є:

- а) формування за якісними ознаками повинні передувати групуванням за кількісними ознаками;
- б) результативні ознаки повинні бути виражені в однакових для всієї системи абсолютних, відносних і середніх показниках;
- в) таблиці системи групувань повинні мати протягом тривалого часу стабільну нумерацію;
- г) інтервали групувань також повинні бути стабільними і без потреби часто не змінюватись.

Дотримання вищенаведених вимог робить систему групувань більш гнучкою і одночасно стабільною, що є необхідною умовою всебічного економіко-статистичного аналізу.



Питання для самоконтролю по темі 3

1. *Поняття про статистичне зведення.*
2. *Завдання і основний зміст зведення.*
3. *Методика статистичного групування.*
4. *Основні правила утворення груп.*
5. *Інтервали статистичних групувань.*
6. *Вибір групувальних ознак.*
7. *Поняття типологічних групувань.*
8. *Поняття структурних групувань.*
9. *Поняття аналітичних групувань.*
10. *Поняття про факторні і результативні ознаки.*
11. *Поняття про вторинні групування.*
12. *Прості і складні групування.*
13. *Поняття про комбіновані групування.*
14. *Багатовимірні групування.*
15. *Основні вимоги до системи групувань.*

ТЕМА 4. СТАТИСТИЧНІ ПОКАЗНИКИ

Питання для теоретичної підготовки

1. Система вимірювачів економічних величин.
2. Види та значення статистичних показників.
3. Абсолютні величини та їх основні види.
4. Відносні величини. Їх значення та основні види.
5. Сутність і значення середніх величин.
6. Види середніх величин і методи їх розрахунку.
7. Основні властивості середньої арифметичної.
8. Структурні середні величини.

1. Система вимірювачів економічних величин

Після проведення зведення і групування статистичних даних та формування статистичної сукупності приступають до *обчислення* узагальнених статистичних показників.



Статистична сукупність – результат статистичного зведення і групування. Це – сукупність об'єктів або явищ суспільного життя, об'єднаних загальним зв'язком, але таких, що розрізняються по ряду варіюючих ознак. Ці об'єкти або явища являють собою елементи (одиниці) статистичної сукупності.

В економічній науці ще не знайдене (і, можливо, ніколи не буде знайдене) єдине мірило, еталон, що був би застосовний до різних економічних величин і дозволяв би вимірювати їх – це головна особливість економічних величин взагалі.

У кожного об'єкта дослідження таких сторін багато, і всякий окремий вимір виявляється неповним, а іноді навіть вводить в оману.



Економічні показники – величини, критерії, рівні, вимірники, які дозволяють судити про стан економіки світу, країни, регіону, підприємства, родини та зміни цього стану, про економічний розвиток, зростання, підйоми або спади.

До теперішнього часу в економічній науці склалися наступні **основні системи вимірників (класифікація) економічних величин**:

1. **Натуральні, трудові й вартісні (грошові)**, хоча принципово можливі й інші (наприклад, енергетичні). Кожна із цих систем має свої особливості, переваги й недоліки та застосовується в різних ситуаціях, що не дозволяє вважати якусь одну з них головною.
2. Прийнято також розрізняти **кількісні** (виражені у фізичних або грошових одиницях – штуках, одиницях ваги, обсягу, довжини, площі, гривнях, доларах) і **якісні** економічні показники, а також **одиничні й групові (агреговані, синтетичні)**.
3. Залежно від області застосування показники розподіляються на **аналітичні, прогнозні, планові, розрахункові, звітні, статистичні** (які застосовуються в статистиці).

Інформація про суспільні явища і процеси створюється, передається і зберігається у вигляді **статистичних показників**. Вони є однією з основних економічних категорій, за допомогою яких відображають кількісну і якісну сторони явищ і процесів.

Адекватність або відповідність показника полягає в його спроможності відобразити ті властивості явищ, які намічені програмою статистичного дослідження. Точність оцінки залежить від структури показника, організації статистичного спостереження та обробки отриманих даних.

2. Види та значення статистичних показників

У господарській практиці вимір економічних величин знаходить своє вираження в **статистичних показниках**, що характеризують властивості економічних об'єктів та необхідні для управління й наукового пізнання. Вони, відповідно, також підрозділяються на натуральні, трудові й вартісні.

Натуральні статистичні показники призначені для зведення багатомірних характеристик споживчих властивостей товарів в одномірні. Приклади: перерахування в умовну продукцію (умовні банки консервів), використання основного параметра (кубічні або квадратні метри при розрахунку тарифів ЖКГ), побудова умовного узагальнюючого показника (різні види палива переводяться у тонни умовного палива – т.у.п.).

Трудові статистичні величини характеризуються показниками трудомісткості продукції, складності праці тощо.

Вартісні статистичні показники – ціна, тариф, маржа, мито тощо.

Узагальнюючі статистичні показники відображають повну сторону досліджуваної сукупності суспільних явищ. Являють собою їхню величину, виражену відповідною одиницею виміру. Ці статистичні величини характеризують обсяги досліджуваних процесів (чисельність працівників, випуск продукції, обсяг товарообігу), їх співвідношення (наприклад, між робітниками та іншими категоріями працівників). У практиці обчислюються різноманітні статистичні показники, що ставляться до багатьох сторін життя суспільства. Різноманіття функцій і цілей, які виконують статистичні показники, визначають їх види.

- *За способом обчислення розрізняють первинні й вторинні (похідні) статистичні показники.*

Первинні статистичні показники отримують в результаті зведення матеріалів статистичного спостереження у формі абсолютних величин (кількість підприємств галузі, вартість основних виробничих

фондів і випущеної ними продукції). На підставі первинних даних обчислюють **вторинні статистичні показники** (середню вартість основних виробничих фондів і випуск продукції в середньому на одне підприємство), а при визначенні інтенсивності використання основних виробничих фондів матимемо похідні показники другого порядку (фондовіддача, фондоємність).

- *За часовою ознакою усі статистичні показники поділяють на інтервальні і моментні.*

Інтервальні статистичні показники характеризують розвиток явищ за окремі періоди часу (наприклад, випуск продукції за тиждень, декаду, місяць, квартал, рік). **Моментні статистичні показники** дають кількісну характеристику стану явища на певну дату, момент часу (наприклад, величину товарних запасів на початок і кінець періоду, чисельність працівників на перше число кожного місяця тощо).

Якщо інтервальні показники можна підсумовувати, то дані, наведені на конкретну дату, підсумовувати недоцільно.

3. Абсолютні величини та їх основні види



Абсолютними величинами в статистиці називаються первинні узагальнюючі показники, які характеризують суспільні явища і процеси в конкретних умовах місця і часу. Абсолютні величини виражають розміри, рівні, обсяги явищ і процесів.

Одержують їх в результаті статистичного спостереження й зведення вихідної інформації. Практично, статистична інформація починає формуватися з абсолютних величин.

Абсолютні величини як узагальнюючі показники характеризують сукупність за її чисельністю (число працівників, кількість підприємств) і

обсягом (валовий випуск продукції, фонд заробітної плати, обсяг роздрібного товарообороту тощо).

Статистика виділяє три види абсолютних величин: *індивідуальні, групові і загальні*.

Індивідуальними називаються такі абсолютні величини, які виражають розміри кількісних ознак окремих одиниць досліджуваної сукупності.

Групові і загальні абсолютні статистичні величини виражають величину ознаки у всіх одиниць даної сукупності або окремих її груп.

Абсолютні величини – завжди числа іменовані, що мають певну розмірність, одиниці виміру.

Одиниці виміру абсолютних величин зводяться до 5-ти типів:

- **Натуральні одиниці** виміру здебільшого відповідають природним або споживчим властивостям предмету, товару й виражаються у фізичних мірах ваги, довжини та ін. (продаж м'яса у кг, тоннах).
- **Умовно-натуральні** одиниці одержують, приводячи різні натуральні одиниці до однієї, прийнятої за основу – еталону.
- **Грошові вартісні одиниці** – як правило, у порівнянних або незмінних цінах для характеристики у вартісному вираженні (наприклад, національний дохід, товарообіг).
- **Трудові одиниці** використовуються для визначення витрат праці на виробництво продукції (людино*години, людино*дні, людино*місяці).
- **Одиниці для виміру часу** (тривалість життя людей вимірюється в роках).

Абсолютні величини є основою для обчислення різних видів відносних і середніх величин, індексів та інших узагальнюючих показників.

4. Відносні величини. Їх значення та основні види

Вивчаючи економічні явища, статистика не може обмежуватися обчисленням тільки абсолютних величин. Абсолютна величина, взята сама по собі, не завжди дає правильну оцінку явища. У багатьох випадках тільки в порівнянні з іншою абсолютною величиною дана величина проявляє свою істинну значимість.



Відносними величинами називають статистичні показники, які виражають кількісні співвідношення між соціально-економічними явищами і процесами.

Їх отримують шляхом порівняння (ділення) двох однойменних, або різнойменних величин.

При розрахунку відносних величин варто мати на увазі, що в чисельнику завжди перебувають показники, що відображають те явище, що вивчається, тобто порівнюваний показник, а в знаменнику – показник, з яким провадяться порівняння, прийнятий за основу або базу порівняння. **База порівняння** виступає як своєрідний вимірник.

Відносні величини показують, у скільки разів порівнювана величина більша (менша) за базисну, або яку частку перша займає в другій, або скільки одиниць однієї величини припадає на одиницю іншої.

В залежності від бази порівняння відносні величини можуть виражатись у формі:

- а) коефіцієнтів – якщо база порівняння приймається за одиницю;
- б) процентів (%) – якщо база порівняння береться за 100;
- в) проміле (‰) – якщо за базу порівняння взято 1000;
- г) продециміле (‱) – якщо база порівняння становить 10 000;
- д) просантиміле (‱‱) – якщо база порівняння прийнята за 100 000.

Існують також іменовані відносні величини (наприклад, показник фондівіддачі: обсяг випуску продукції, середньорічна вартість основних фондів).

Розрахунок відносних величин може бути правильним лише за умови, що показники, які рівняються, є порівнянними.

За своїм пізнавальним значенням відносні величини підрозділяються на наступні види:

- відсоток виконання плану;
- відносні величини динаміки:
 - ланцюгові;
 - базисні.
- відносні величини структури (частки або питомої ваги);
- відносні величини координації;
- відносні величини інтенсивності;
- відносні величини рівня економічного розвитку;
- відносні величини порівняння.

Відсоток виконання плану

Розраховується шляхом поділу фактичного рівня досліджуваного показника на його плановий рівень.

$$\frac{\text{Факт}}{\text{План}} \times 100 \% , \quad (4.1)$$

Відносні величини динаміки

Характеризують зміну досліджуваного явища в часі. Виявляють напрямок розвитку й вимірюють його інтенсивність. Розрахунок відносних величин виконується у вигляді темпів росту та інших показників динаміки. Розраховуються базисні й ланцюгові показники.

Ланцюгові величини – величини зі змінною базою порівняння. Відношення рівня кожного наступного періоду до попереднього (ВВДЛ)

$$\hat{A}\hat{A}\ddot{E} = \frac{\acute{O}_i}{\acute{O}_{i-1}}, \quad (4.2)$$

Базисні – відношення рівня кожного наступного періоду до одного, прийнятого за базу порівняння (ВВДБ).

$$\hat{A}\hat{A}\ddot{A} = \frac{\acute{O}_i}{\acute{O}_0}, \quad (4.3)$$

Відносні величини структури

Характеризують склад досліджуваних сукупностей. Вони розраховуються, як відношення абсолютних величин *кожного* з елементів сукупності до абсолютних величин всієї сукупності, тобто як відношення частини до цілого, і являє собою питому вагу частини в цілому. Як правило, відносні величини структури виражаються у відсотках (база порівняння приймається за 100 %).

$$\hat{A}\hat{A}_{\tilde{n}\ddot{o}\acute{o}\acute{o}\acute{o}\acute{o}\acute{o}} = \frac{\div\grave{a}\tilde{n}\ddot{o}\acute{o}\acute{o}}{\acute{o}^3\grave{e}\grave{a}} \times 100\%, \quad (4.4)$$

Відносні величини координації

Являє собою одну з різновидів показників порівняння. Вони застосовуються для статистичної сукупності. Показують, у скільки разів порівнювана частина сукупності більше або менше частини, що приймається за основу або базу порівняння

$$\hat{A}\hat{A}_{\grave{e}\tilde{n}\ddot{o}\acute{o}\acute{o}\acute{o}\acute{o}\acute{o}}_i = \frac{\div\grave{a}\tilde{n}\ddot{o}\acute{o}\acute{o}}{\div\grave{a}\tilde{n}\ddot{o}\acute{o}\acute{o} \quad \acute{a}\grave{a}\grave{c}\acute{i}\grave{a}\grave{a}}, \quad (4.5)$$

Відносні величини інтенсивності

Показують, наскільки поширене досліджуване явище в тому або іншому середовищі. Вони характеризують співвідношення різнойменних, незв'язаних між собою абсолютних величин.

Розраховується поділом абсолютних величин досліджуваного явища на абсолютні величини, що характеризують обсяг середовища, в

[illegible]

Це показники, що характеризують розміри виробництва різних видів продукції на душу населення.

$$\hat{A}\hat{A}^{\dagger} \frac{\partial^3 \hat{a} \hat{f} \hat{y}}{\partial \hat{a} \hat{e} \hat{t} \hat{t} \hat{t}} \cdot \frac{\partial \hat{f} \hat{\zeta} \hat{a} \hat{e} \hat{o} \hat{e} \hat{o}}{\partial \hat{e} \hat{n} \hat{a} \hat{e} \hat{u} \hat{t}^3 \hat{n} \hat{o} \hat{u} \hat{t} \hat{a} \hat{n} \hat{d} \hat{e} \hat{a} \hat{f} \hat{y}} = \frac{\partial^3 \hat{f} \hat{e} \hat{e} \hat{t} \hat{a} \hat{n} \hat{y} \hat{a} \hat{a} \hat{e} \hat{o} \hat{t} \hat{a} \hat{t} \hat{o} \hat{o} \hat{a} \hat{a} \hat{t} \hat{o} \hat{a} \hat{e} \hat{o}^3 \hat{\zeta}}{\partial \hat{e} \hat{n} \hat{a} \hat{e} \hat{u} \hat{t}^3 \hat{n} \hat{o} \hat{u} \hat{t} \hat{a} \hat{n} \hat{d} \hat{e} \hat{a} \hat{f} \hat{y}}, \quad (4.7)$$

Характеризують кількісне співвідношення однойменних показників, але стосовно до різних об'єктів статистичного спостереження (наприклад, національний дохід Польщі зіставляється з національним доходом України).

Для зведеної кількісної характеристики багатьох явищ і процесів суспільного життя статистика в якості одного з методів економічного аналізу (як 3-й етап статистичного дослідження) широко використовує такий розповсюджений узагальнюючий показник як середня величина (середній відсоток виконання плану, середня частка тощо). Вона дає узагальнюючу характеристику однорідних елементів масових явищ, які мають різне кількісне значення (варіацію) в залежності від конкретних умов.

58



Середня величина (\bar{X}) – це узагальнюючий показник, який характеризує однорідну сукупність явищ за якою-небудь кількісною варіаційною ознакою в даних умовах місця і часу.

Тільки за допомогою середньої величини можна охарактеризувати сукупність за кількісною варіаційною ознакою.

Середня величина – це узагальнююча характеристика сукупності, що має три основні, фундаментальні властивості:

1. середня відображається єдиним числом;
2. середня розташована між найбільшим і найменшим варіантом сукупності або дорівнює їм;
3. середня – це наукова знакова модель досліджуваної сукупності.

Середні величини використовують для порівняння показників двох і більше об'єктів. Середніми величинами користуються для характеристики зміни рівнів явищ в часі.

До середніх звертаються при вивченні взаємозв'язків між явищами та їх ознаками. Середні величини застосовують для проведення факторного аналізу явищ, з метою виявлення невикористаних резервів.

Багатогранність суспільних явищ обумовлює виняткову важливість застосування середніх величин в економіко-статистичних дослідженнях. Вони є активним засобом управління, планування і прогнозування економіки держави.



Середні величини – це узагальнюючі статистичні показники, в яких знаходять своє вираження дія загальних умов, закономірність досліджуваного явища. Статистичні середні величини розраховується на підставі масових даних коректно статистично організованого масового спостереження (суцільного або вибіркового).

Статистична середня буде об'єктивна й типова, якщо вона розраховується за масовим даними для якісно однорідної сукупності.

Середня величина є величина абстрактна, тому що характеризує значення абстрактної одиниці, і тому відрізняється від структури сукупності. У середній величині як і у всякій абстракції здійснюється діалектична єдність окремого й загального. Середня величина відображає те загальне, що складається в кожному окремому одиничному об'єкті, завдяки цьому середня одержує велике значення для виявлення закономірностей, які властиві масовим суспільним явищам і непритаманні одиничним явищам.

Середня величина є відбиттям значень досліджуваної ознаки й вимірюється в тій же розмірності, що й ця ознака. Сполучення загальних середніх із груповими дає можливість виявити та обмежити якісно однорідні сукупності.

6. Види середніх величин і методи їх розрахунку

У практиці статистичної обробки матеріалу виникають різні завдання, є особливості досліджуваних явищ і тому для їх рішення потрібні різні відомості.

В статистиці найчастіше використовують наступні види середніх величин:

- середня арифметична (проста й зважена);
- середня гармонійна (проста й зважена);
- середня геометрична;
- середня квадратична (проста й зважена);
- середня хронологічна;
- мода;
- медіана.

Питання про те, який вид середньої необхідно застосувати в окремому випадку, вирішується шляхом окремого логічного аналізу досліджуваної сукупності, визначається змістовним значенням досліджуваного

явища. Тільки тоді середню величину застосовано правильно, коли одержані величини мають логіку та реальний економічний зміст.

Введемо наступні поняття й позначення.



Середня величина – кількісна характеристика ознаки в статистичній сукупності. Вона визначає величину ознаки, віднесену до одиниці сукупності й абстраговану від індивідуальних особливостей окремих одиниць. **Середня величина** розраховується на одиницю статистичної сукупності, як співвідношення обсягу явища (загального розміру ознаки) і обсягу сукупності (загального числа одиниць).

Ознака, по якій вираховують середню – усереднюєма **ознака** X_i .

Величина X_i у кожній одиниці сукупності називається **індивідуальним значенням (варіантом)** $X_1, X_2, X_3, \dots, X_n$.

Частота - повторюваність індивідуальних значень ознаки – f .

Середня арифметична

Найбільш поширеним видом середніх величин в статистці є середня арифметична. Вона розраховується в тих випадках, коли обсяг X визначається як сума його значень в окремих одиниць досліджуваної статистичної сукупності. Залежно від характеру вихідних даних середня арифметична визначається так:

1) як середня арифметична незважена (**проста**): розподілом кількості зведеної ознаки на число показань:

$$\overline{X}_{\text{ад.їд.}} = \frac{\sum X_i}{n} = \frac{X_1 + X_2 + \dots + X_n}{n}, \quad (4.8)$$

2) або як середня арифметична **зважена**:

$$\overline{X}_{\text{ад.џâ}} = \frac{\sum X_i f_i}{\sum f_i} = \frac{X_1 f_1 + X_2 f_2 + \dots + X_n f_n}{f_1 + f_2 + \dots + f_n}, \quad (4.9)$$

Часто доводиться розраховувати середнє значення ознаки по ряду розподілу, коли теж саме значення ознаки зустрічається кілька разів. Частоти окремих варіантів можуть бути виражені не тільки абсолютними величинами, але й відносними величинами – **частотями** (w_i).

Замінивши абсолютні значення частот відповідними відносними величинами, отримаємо:

$$X_{\text{середнє}} = \frac{\sum X_i w_i}{\sum w_i}, \quad (4.10)$$

Зважена середня враховує різні значення окремих варіантів у межах сукупності.

Часто обчислення середніх величин доводиться проводити за даними, згрупованими у вигляді інтервальних рядів розподілу. Тоді варіанти ознаки, з яких обчислюється середня, представлені у вигляді інтервалу (від і до). Для обчислення середньої величини необхідно у кожному варіанті визначити **серединне значення** X^1 , після чого зробити зважування звичайним порядком:

$$X^1 * f, \quad (4.11)$$

У закритому інтервалі серединне значення визначається як напівсума значень нижньої й верхньої границь. Іноді завдання обчислення середньої по величинах інтервального ряду ускладнюються тим, що невідомі крайні границі початкового й кінцевого інтервалів. У цьому випадку передбачається, що відстань між межами даного інтервалу така ж, як і в сусідньому інтервалі. Цей прийом заснований на припущенні, що окремі конкретні варіанти рівномірно розподілені усередині інтервалу.

Середня гармонічна

У статистичній практиці часто зустрічаються випадки, коли середню величину потрібно обчислювати за формулою середньої гармонічної. Це відбувається тоді, коли підсумовуванню підлягають не самі варіанти, а обернені їм числа, тобто статистична інформація не містить частот по

окремих варіантах сукупності, а представлена як їхній добуток ($X_i * f$). У цьому випадку застосовують формулу середньої гармонічної зваженої:

$$X_{\text{середньої гармонічної зваженої}} = \frac{\sum M_i}{\sum \frac{M_i}{X_i}} = \frac{M_1 + M_2 + \dots + M_n}{\frac{M_1}{X_1} + \frac{M_2}{X_2} + \dots + \frac{M_n}{X_n}}, \quad (4.11)$$

де $M_i = X_i * f_i$, тобто X_i і f_i перебувають не у вільному стані, а у зв'язаному.

Середню гармонічну зважену застосовують у тих випадках, коли є дані про індивідуальні значення ознаки у загальній сукупності і загальний обсяг сукупності, але в готовому виді немає частот.

У тому випадку, якщо обсяги явищ, тобто добутку по кожній ознаці, рівні – застосовується середня гармонічна проста.

$$\overline{X}_{\text{середньої гармонічної простої}} = \frac{n}{\sum \frac{1}{X_i}}, \quad (4.12)$$

де $\sum \frac{1}{X_i}$ – сума зворотних значень, варіант;

n – число варіантів.

Середня геометрична

Середній темп росту обчислюється як середня геометрична темпів росту за окремі відрізки часу або як корінь, ступінь якого визначається числом періодів, а під коренем береться темп росту за весь період.

Це величина, використовується як середня з відносин або в рядах розподілу, представлених у вигляді геометричної прогресії.

$$\overline{X}_{\text{середньої геометричної}} = \sqrt[n]{\prod (X_i)} = \sqrt[n]{X_1 \cdot X_2 \cdot X_3 \cdot \dots \cdot X_n}, \quad (4.13)$$

Цією середньою зручно користуватися, коли приділяється увага не абсолютним різницям, а відноsinам 2-х чисел, тож середня геометрична використовується в розрахунках середніх темпів динаміки.

Середня квадратична

Використовується в тих випадках, коли у сукупність включаються одиниці квадратної функції – при обчисленні показників варіації.

Середня квадратична проста

$$\overline{X}_{\text{кв.пр.}} = \sqrt{\frac{\sum X_i^2}{n}}, \quad (4.14)$$

Середня квадратична зважена

$$\overline{X}_{\text{кв.зв.}} = \sqrt{\frac{\sum X_i^2 \cdot f_i}{\sum f_i}}, \quad (4.15)$$

Середня хронологічна

$$\overline{X}_{\text{хр.}} = \frac{\frac{1}{2}X_1 + X_2 + X_3 + \dots + \frac{1}{2}X_n}{n-1}, \quad (4.16)$$

Використовується в тому випадку, коли показники представлені на певний момент часу й у вигляді рівних інтервалів (середній рівень варіантів моментного ряду).

7. Основні властивості середньої арифметичної

Середня арифметична має деякі математичні властивості, що мають практичне значення для спрощеного обчислення середньої за даними варіаційного ряду, які ми наводимо без доказів.

Найважливіші з них такі:

1. Від зменшення або збільшення частот кожного значення ознаки X_i в n разів величина середньої арифметичної не зміниться. Якщо всі частоти розділити або помножити на будь-яке число d , то величина середньої не зміниться. Ця властивість дає можливість замінити частоти питомими вагами – **частотями**. Також якщо обчислювати середнє за формулою *середньої арифметичної простої*, тоді середнє обчислюється так:

$$\text{якщо } d \text{ у відсотках (\%)}, \quad \bar{X} = \frac{\sum X_i * d}{100}, \quad (4.17)$$

$$\text{якщо } d \text{ у частках одиниці}, \quad \bar{X} = \sum X_i * d, \quad (4.18)$$

2. Загальний множник індивідуальних значень ознаки може бути винесений за знак середньої,

$$\overline{KX} = K * \bar{X}, \quad (4.19)$$

3. Середня суми (різниці) двох або декількох величин дорівнює сумі (різниці) їх середніх:

$$\overline{X \pm Y} = \bar{X} \pm \bar{Y}, \quad (4.20)$$

4. Якщо $X_i = C$, де $C = \text{const}$, тоді $\bar{X} = \bar{C} = C$

5. Сума відхилень значень ознаки X_i від їх середньої завжди дорівнює нулю:

$$\sum (X_i - \bar{X}) = 0$$

Викладені вище властивості середньої арифметичної дозволяють у багатьох випадках спростити її розрахунок: можна з усіх значень ознаки відняти довільну постійну величину та різницю скоротити на загальний множник. А потім обчислену середню помножити на загальний множник і додати довільну постійну величину.

Формула середньої арифметичної зваженої отримає наступний вигляд:

$$\overline{X} = m_1 * i + A, \quad (4.21)$$

де $m_1 = \frac{\sum (\frac{X - A}{i}) * f}{\sum f}$ – момент першого порядку;

X – варіант;

f – частота;

A – довільно взяте число;

i – розмір інтервалу.

Середня m_1 від значення $\frac{X - A}{i}$ називається **моментом першого порядку**, а спосіб обчислення середньої – **методом відліку від умовного нуля**.

8. Структурні середні величини

Для характеристики структури сукупності застосовуються особливі показники, до яких відносять моду і медіану і які можна назвати **структурними середніми**.

Мода (M_0) – це варіант, що зустрічається найчастіше, або ж модою називається те значення ознаки, що відповідає точці **max** теоретичної кривої ряду розподілу. Мода являє собою найбільш типове значення.

У дискретному ряді мода – це варіанта з найбільшою частотою.

В інтервальному варіаційному ряді моду вважають центральним варіантом так званого модального інтервалу, тобто того інтервалу, що має найбільшу частоту (частковість).

Конкретне значення моди для інтервального ряду визначається формулою

$$M_0 = Xm_0 + Im_0 * \frac{(fm_0 - fm_{0-1})}{(fm_0 - fm_{0-1}) + (fm_0 - fm_{0+1})}, \quad (4.22)$$

де Xm_0 – нижня межа модального інтервалу;

Im_0 – величина модального інтервалу;

fm_0 – частота, що відповідає модальному інтервалу;

fm_{0-1} – частота, що передуює модальному інтервалу;

fm_{0+1} – частота інтервалу, що впливає за модальним.

Медіана (M_e) – величина, що ділить чисельність упорядкованого варіаційного ряду на дві рівні частини: одна частина має значення ознак, що варіюють, менші ніж середній варіант, а інша – більші.

Якщо припустити, що всередині зростання або зниження досліджуваної ознаки відбувається рівномірно, то формула медіани в інтервальному ряді розподілу буде мати такий вигляд:

$$M_e = X_{me} + i * \frac{\sum \frac{f}{2} - S_{me-1}}{f_{me}}, \quad (4.23)$$

де X_{me} – нижня межа медіанного інтервалу;

i – розмір медіанного інтервалу;

$\sum f$ – сума частот;

$\sum \frac{f}{2}$ – порядковий номер медіани;

S_{me-1} – сума накопичених частот, що передують медіанному інтервалу;

f_{me} – частота медіанного інтервалу.

Медіана знаходить практичне застосування внаслідок особливої властивості: сума абсолютних відхилень членів ряду від медіани є величина найменша:

$$\sum (X_i - M_e) = \min, \quad (4.24)$$



Питання для самоконтролю по темі 4

1. *Поняття про економічні показники.*
2. *Визначення статистичних показників.*
3. *Способи вирахування статистичних показників.*
4. *Одиниці виміру статистичних показників.*
5. *Система вимірювачів статистичних показників.*
6. *Первинні та вторинні статистичні показники.*
7. *Поняття про абсолютні величини.*
8. *Поняття про відносні величини.*
9. *Види відносних величин.*
10. *Форми представлення відносних величин.*
11. *Однчасне використання абсолютних і відносних величин.*
12. *Сутність середніх величин.*
13. *Значення середніх величин.*
14. *Види середніх величин.*
15. *Способи обчислення середніх величин.*
16. *Властивості середньої арифметичної.*
17. *Поняття про структурні середні.*

ТЕМА 5. АНАЛІЗ РЯДІВ РОЗПОДІЛУ

Питання для теоретичної підготовки

1. Сутність і характеристики варіації.
2. Показники варіації та способи їх розрахунку.
3. Види дисперсій і правило «Додавання дисперсій». Дисперсія альтернативної (якісної) ознаки.
4. Поняття рядів розподілу, їх види.
5. Форми рядів розподілу, та їх характеристика.
6. Графічне зображення рядів розподілу.
7. Сутність і характеристики варіації.

Різні кількісні величини того ж самого параметру ознаки у різних об'єктів всередині досліджуваної сукупності в статистиці називають **варіантами** (позначаються X_i), а саме розходження індивідуальних значень ознаки всередині досліджуваної сукупності називається **варіацією ознаки**.

Термін «**варіація**» походить від лат. «**variatio**» – зміна, коливання, розходження. Варіація обумовлена впливом на ознаку різних факторів, їх сполученням у кожному окремому випадку.

Колівання окремих значень характеризують показники варіації. Однак не всі розходження прийнято називати варіацією.



Під варіацією в статистиці розуміють такі кількісні зміни величини досліджуваної ознаки в межах однорідної сукупності, які обумовлені перехресним впливом дії різних факторів та їх сполученням.

Розрізняють **випадкову** й **систематичну** варіації. Аналіз систематичної варіації дозволяє оцінити ступінь залежності змін у досліджуваній ознаці від визначальних її факторів. Наприклад, можна оцінити, наскільки

однорідною є сукупність. Для характеристики сукупностей і обчислених середніх величин важливо знати, яка варіація досліджуваної ознаки ховається за середніми.



Варіація, тобто коливання, мінливість значень будь-якої ознаки є властивістю статистичної сукупності. Вона зумовлена дією безлічі взаємопов'язаних причин, серед яких є основні й другорядні. Основні причини формують *центр розподілу*, другорядні – *варіацію ознак*, сукупна їх дія – *форму розподілу*.

Наприклад, прибуток підприємства залежить від продуктивності праці, якості продукції, що випускається, кон'юнктури ринку збуту та інших об'єктивних і суб'єктивних факторів. Сумісна дія їх і різне поєднання зумовлюють той чи інший рівень рентабельності окремих підприємств, а також закономірність розподілу підприємств за цією ознакою.

Статистичні характеристики центру розподілу (середня, мода, медіана) відіграють важливу роль у вивченні статистичних сукупностей. Однак середні величини, характеризуючи варіаційний ряд одним числом, не відображають мінливості значень ознаки, що спостерігалися, тобто їх відхилення від середнього або варіацію. В одних сукупностях індивідуальні значення ознаки значно відхиляються від центру розподілу, в інших – тісно групуються навколо нього, а відтак виникає потреба оцінити поряд з характеристиками центру розподілу міру і ступінь варіації. Чим менша варіація, тим однорідніша сукупність, отже тим більш надійні й типові характеристики центру розподілу, насамперед середні величини.

2. Показники варіації і способи їх розрахунку

Для виміру і оцінки варіації використовують систему абсолютних і відносних характеристик – **показників варіації**. Показники варіації і середні величини дозволяють судити про характерні риси досліджуваних явищ, представлених у вигляді варіаційних рядів.

У статистиці для характеристики розсіювання (коливання) ознак використовуються наступна система показників варіації :

■ **абсолютні характеристики варіації:**

- розмах варіації;
- середнє лінійне відхилення (просте й зважене);
- дисперсія (проста й зважена);
- середнє квадратичне відхилення(просте, зважене).

■ **відносні характеристики варіації (коефіцієнти варіації):**

- відносне лінійне відхилення (лінійний коефіцієнт варіації);
- коефіцієнт варіації (квадратичний);
- коефіцієнт осциляції.

Абсолютні характеристики варіації – величини іменовані і мають одиниці виміру варіюючої ознаки. При порівнянні варіації різних ознак використовуються відносні характеристики – коефіцієнти варіації, які обчислюють як відношення абсолютних характеристик варіації до характеристики центру розподілу.

Абсолютні показники варіації

Найпростішим показником варіації є **розмах варіації (R)**, що дорівнює різниці між найбільшим (X_{\max}) і найменшим (X_{\min}) варіантами.

$$R = X_{\max} - X_{\min} , \quad (5.1)$$

Це найпростіший і досить наближений показник варіації, тому що фіксує тільки крайні (як правило нехарактерні) значення й не відображає усіх відхилень у досліджуваному ряді. Необхідний тільки для окреслення кола завдань.

Більший інтерес представляють міри розсіювання результатів спостережень навколо середніх величин.

Відхилення варіантів від середньої арифметичної характеризують **різниці** ($X_i - \bar{X}$)

де X_i – значення показника;

\bar{X} – середнє арифметичне значення.

Але проста сума відхилень (проста або зважена) не може характеризувати варіацію навколо середньої, тому що за 5-ю властивістю середньої дорівнює нулю. Тому в різницях беруть або їх абсолютні величини, або їх квадрати. Відповідно до цього одержують різні показники варіації, що враховують розсіювання усіх одиниць досліджуваної сукупності навколо середньої.

1. **Середнє лінійне відхилення** (\bar{l}), визначається як середнє арифметичне з відхилень індивідуальних значень варіантів від середньої арифметичної без урахування знаків цих відхилень:

$$\bar{l}_{i\bar{o}} = \frac{\sum |X_i - \bar{X}|}{n}, \quad (5.2)$$

або

$$\bar{l}_{\zeta\hat{a}} = \frac{\sum |X_i - \bar{X}| * f}{\sum f}, \quad (5.3)$$

2. На практиці міру варіації більш об'єктивно відображає показник **дисперсії** (σ^2) – середня арифметична квадратів відхилень варіантів від їх середньої арифметичної $(X_s - \bar{X})^2$, у найпростішому випадку незгрупованих даних:

$$\sigma_{i\bar{o}}^2 = \frac{\sum (X_s - \bar{X})^2}{n}, \quad (5.4)$$

або дисперсія для згрупованих даних і для інтервальних рядів:

$$\sigma_{\text{сг.}}^2 = \frac{\sum (X_s - \bar{X})^2 * f}{\sum f}, \quad (5.5)$$

де $\sum f = n$.

3. Міра розсіювання повинна бути виражена в тих же одиницях, що і значення ознаки, тому замість дисперсії як показник варіації частіше використовують корінь квадратний з дисперсії.

Арифметичне значення кореня квадратного з дисперсії називається **середнім квадратичним відхиленням**, або **стандартним відхиленням**, і позначається σ .

$$\sigma = \sqrt{\sigma^2} \quad (\text{просте або зважене}), \quad (5.6)$$

На відміну від дисперсії, цей показник, що також показує ступінь варіації ознаки, має розмірність самої ознаки, а не його квадрата, що представляє певну зручність.

Далі ми побачимо, що стандартне відхилення має важливе значення в теорії оцінювання невідомих параметрів (наприклад, середнього генеральної сукупності) і в теорії помилок вибіркового спостереження.

Середнє квадратичне відхилення є мірилом надійності середньої: чим менше середнє квадратичне відхилення тим повніше середнє арифметичне відображає властивості всієї представленої сукупності.

Дисперсія (як і інші статистичні характеристики) володіє рядом властивостей, які дозволяють спростити її обчислення.

1. Дисперсія постійної величини дорівнює нулю.
2. Якщо усі значення варіантів зменшити (або збільшити) на те саме число А, то дисперсія від цього не зміниться:

$$\sigma_{X_i - A}^2 = \sigma_{X_i}^2, \quad (5.7)$$

3. Якщо всі значення варіантів зменшити (або збільшити) у B разів, то дисперсія зменшиться (або збільшиться) від цього у B^2 разів. Тож, як слідство, середнє квадратичне відхилення зменшиться (або збільшиться) у B разів:

$$\sigma_{(X_i * B)}^2 = \sigma^2 * B^2; \quad \sigma_{\frac{X_i}{B}} = \frac{\sigma}{B}, \quad (5.8)$$

4. Якщо обчислити дисперсію від будь-якої величини A , що у тім або іншому ступені відрізняється від середньої арифметичної (\bar{X}), то така дисперсія завжди буде більше дисперсії, обчисленої від середньої арифметичної:

$$\sigma_A^2 > \sigma_X^2, \quad (5.9)$$

при цьому більше на цілком певну величину – на квадрат різниці між середньою й цією умовно взятою величиною, тобто на $(\bar{X} - A)^2$.

$$\sigma^2 = \sigma_A^2 + (\bar{X} - A)^2, \quad (5.10)$$

$$\sigma^2 = \sigma_A^2 - (\bar{X} - A)^2, \quad (5.11)$$

Відносні показники варіації

Показники відносного розсіювання визначають шляхом розподілу міри абсолютного розсіювання на характеристики центру розподілу (найчастіше – середню арифметичну) й виражаються у відсотках. До таких показників належать:

1. **Відносне лінійне відхилення (лінійний коефіцієнт варіації)**, який дорівнює вираженому у відсотках відношенню середнього лінійного відхилення до середнього арифметичного :

$$V_{\bar{l}} = (\bar{l} / \bar{X}) * 100\% , \quad (5.12)$$

де \bar{l} – середнє лінійне відхилення;

\bar{X} – середня величина.

2. Для характеристики того, наскільки середня арифметична добре представляє варіаційний ряд, використовують **коефіцієнт варіації (квадратичний)** рівний вираженому у відсотках відношенню середнього квадратичного відхилення до середнього арифметичного:

$$V_{\sigma} = (\sigma / \bar{X}) * 100\% , \quad (5.13)$$

де σ – середнє квадратичне відхилення;

\bar{X} – середня величина.

Порівняння середніх квадратичних відхилень, обчислених для однотипних варіаційних рядів, дозволяє встановити, де розсіювання значень ознак навколо середньої арифметичної більше.

Якщо варіаційні ряди мають однакові середні арифметичні, то середня арифметична ряду з меншим коефіцієнтом варіації – більше представницька. Очевидним є й наступне твердження: якщо варіаційні ряди мають однакові середні квадратичні відхилення, то більш представницькою є середня арифметична ряду, в якого коефіцієнт варіації менший.

Коефіцієнти варіації є найпоширенішими показниками коливання, які використовуються для оцінки типовості середніх величин і однорідності досліджуваних сукупностей.

При цьому виходять із такого співвідношення: якщо $V < 30\%$, це свідчить про велику однорідність ознак досліджуваної сукупності; якщо

30 % < \mathcal{V} < 100 %, то однорідність ознак досліджуваної сукупності – задовільна; а якщо $\mathcal{V} > 100$ %, тоді можна зробити висновок, що досліджувана сукупність неоднорідна.

3. **Коефіцієнт осциляції** – визначається як відношення розмаху варіації до середньої величини ознаки й характеризує відносну неуважність або коливання крайніх значень ознаки навколо середньої арифметичної:

$$\mathcal{V}_R = (R / \overline{X}) , \quad (5.14)$$

де R – розмах варіації;

\overline{X} – середня величина.

3. Види дисперсій і правило «Додавання дисперсій»

У загальному випадку варіація результативної ознаки обумовлена різними факторами в їх сукупності, а не тільки впливом одного з них. Якщо статистичну сукупність розбити на групи за будь-якою ознакою, то поряд з вивченням варіації результативної ознаки по всій сукупності в цілому під впливом усіх факторів одержуємо можливість вивчити варіацію для кожної зі складових всієї сукупності груп окремо. Можна також вивчити при цьому варіацію між групами.

У найпростішому випадку всю вихідну сукупність розбивають на окремі групи за одним фактором. Тоді наведений вище аналіз варіації зводиться до розрахунку й аналізу трьох видів дисперсії: загальної, внутрішньогрупової й міжгрупової.

Загальна дисперсія вимірює варіацію результативної ознаки по всій сукупності під впливом всіх факторів, що обумовили цю варіацію.

Міжгрупова дисперсія δ^2 характеризує систематичну варіацію під впливом ознак – факторів, покладених в підставу групування. Вона

дорівнює середньому квадрату відхилень групових (часток) середніх від загальної середньої для всієї сукупності:

$$\delta^2 = \frac{\sum (X_i - \bar{X})^2 * f}{\sum f}, \quad (5.15)$$

де f – чисельність одиниць у групі (частота).

Внутрішньогрупова дисперсія є вже відома нам дисперсія (для усієї сукупності, яка називється загальною), але тепер ця формула застосовується тільки до окремої групи. Відповідно й позначається вона σ^2 , але вже з індексом i , що підкреслює, що розрахунок виконується для окремої i -тої групи, тобто σ_i^2 .

Внутрішньогрупова дисперсія відображає випадкову варіацію, тобто ту її частину, що обумовлена впливом інших (неврахованих) факторів, відмінних від підстави угруповання. По окремих внутрішньогрупових дисперсіях, розглядаючи їх як значення деякої особливої ознаки, розраховують середню по внутрішньогрупових дисперсіях, що вже характеризує варіацію по усій сукупності в цілому під впливом усіх інших (неврахованих) факторів, відмінних від підстави групування.

Існує проста й важлива формула, що зв'язує загальну дисперсію, міжгрупову дисперсію й середню по внутрішньогрупових дисперсіях:

$$\sigma^2 = \delta^2 + \sigma_i^2, \quad (5.16)$$

Це означає, що загальна дисперсія дорівнює сумі міжгрупової дисперсії й середньої по внутрішньогрупових дисперсіях. Отже, знаючи дві з трьох дисперсій, можна завжди знайти й третю.

Правило додавання дисперсій показує, що чим більше частка міжгрупової дисперсії в загальній дисперсії, тим сильніше вплив груповальної ознаки на досліджувану результативну ознаку. Такі міркування природно приводять до кількісної характеристики такого впливу, мірі

стохастичного зв'язку між ознаками. Вона називається **емпіричним коефіцієнтом детермінації** і позначається η^2 , характеризуючи чинність впливу групувальної ознаки на утворення загальної варіації:

$$\eta^2 = \frac{\delta^2}{\sigma^2}, \quad (5.17)$$

При відсутності зв'язку він просто дорівнює нулю, при чисто функціональному зв'язку – 1. У загальному випадку коефіцієнт детермінації приймає значення між « 0 » і « 1 ». Це видно й з правила додавання дисперсій.

Крім коефіцієнта детермінації використовують також **емпіричне кореляційне відношення**, що являє собою корінь квадратний з коефіцієнта детермінації. І знову воно достатньо підходить для виміру лінійного зв'язку.

У загальному випадку нелінійного зв'язку переважніше використовують, що вірніше, коефіцієнт детермінації. Якщо зв'язок відсутній, то кореляційне відношення дорівнює нулю й усі групові середні рівні між собою, а міжгрупової варіації в цьому випадку немає. Групувальна ознака при цьому ніяк не впливає на утворення загальної варіації. Якщо зв'язок функціональний, то кореляційне відношення дорівнює 1. Дисперсія групових середніх дорівнює загальній дисперсії і міжгруповій дисперсії, тому внутрігрупової варіації не буде. Таким чином, групувальна ознака повністю визначає варіацію досліджуваної результативної ознаки.

Дисперсія альтернативної (якісної) ознаки

У статистиці поряд з показниками варіації кількісної ознаки визначають показники варіації альтернативної ознаки. **Альтернативними** є ознаки, якими володіють одні одиниці досліджуваної сукупності й не володіють інші (наприклад, стать: у цьому випадку це взаємовиключні варіанти).

При статистичному вираженні коливання альтернативних ознак, наявність досліджуваної ознаки позначається « 1 », а її відсутність « 0 ». Частка

варіантів, що володіють досліджуваною ознакою, позначається p , а частка варіантів, що не володіють ознакою, – q . При цьому:

$$p + q = 1$$

Їх середнє значення дорівнює:

$$\bar{X} = \frac{\sum X_i * f}{\sum f} = \frac{1 * p + 0 * q}{p + q} = p, \quad (5.18)$$

тобто частка одиниць, що володіють досліджуваною ознакою дорівнює p .

Дисперсія альтернативної ознаки дорівнює добутку частки одиниць, що володіють ознакою, й частки одиниць, що не володіють їю.

$$\sigma_p^2 = p * q, \quad (5.19)$$

$$\sigma_p^2 = \frac{\sum (X_i - \bar{X})^2 * f}{\sum f} = \frac{(1-p)^2 * p + (0-p)^2 * q}{p+q} = p * q, \quad (5.20)$$

4. Поняття рядів розподілу, їх види



В результаті статистичного групування отримують ряди цифрових показників, які характеризують розподіл одиниць досліджуваної сукупності за варіаційною ознакою. Такі ряди називаються **рядами розподілу**.

Ряд розподілу складається з двох елементів – варіантів та частот. **Варіанти** « X » – це окремі значення групувальної ознаки, які розташовані у певній послідовності. **Частоти** « f » – це числа, які показують, скільки разів певне значення ознаки зустрічається у сукупності, або скільки одиниць сукупності припадає на кожну групу.

Ряди розподілу відіграють важливу роль при вивченні складу сукупності за досліджуваною ознакою, закономірностей розподілу, використовуються при визначенні середніх величин, показників варіації, взаємозв'язку тощо.

Залежно від характеру групувальної ознаки ряди розподілу поділяються на **атрибутивні** та **варіаційні (кількісні)**.

В атрибутивних рядах розподілу варіанти не мають чисельного вираження.

Варіаційні ряди, в яких варіанти мають числовий вираз, поділяються на дискретні та інтервальні. У дискретних рядах варіанти являють собою дискретні числа, в інтервальних – це інтервали групування.

У тому випадку, коли виконують групування за двома й більше ознаками, отримують комбінаційний ряд розподілу.

При побудові атрибутивних рядів розподілу варіанти потрібно розташувати в логічній послідовності. При використанні дискретних та інтервальних варіаційних рядів варіанти записують в порядку зростання. Для інтервальних рядів важливим є чітке розмежування варіант.

Розрізняють ряди розподілу з абсолютними, відносними та нагромадженими частотами. Нагромаджені частоти ще називають кумулятивними. У першому випадку частоти являють собою абсолютні числа, у другому – питому вагу або частку кожної групи.



Ряди розподілу з абсолютними частотами характеризують *склад сукупності*, а з відносними – *структуру сукупності*.



Кумулятивні частоти знаходять шляхом їх сумування по групах.

Щільність розподілу – це кількість одиниць сукупності, що припадає на одиницю ширини інтервалу групувальної ознаки. Розрізняють **абсолютну** щільність і **відносну**:

$$f_d = \frac{f}{i} ; \quad f = \frac{p}{i} , \quad (5.21)$$

де f – частота;
 p – частка (доля);
 i – величина (розмір) інтервалу.

Інтерполяція у рядах розподілу визначає, скільки одиниць сукупності (або процентів) мають значення ознаки менше за завдане. Для інтерполяції використовують як абсолютні так і відносні нагромаджені частоти.

За допомогою інтерполяції визначають, менше якого значення мають ознаку певна кількість (питома вага) одиниць сукупності.

4. Форми рядів розподілу та їх характеристика

Різноманітність статистичних сукупностей – передумова різних форм співвідношення частот і значень варіюючої ознаки. За своєю формою розподіли поділяють на такі види: одно-, дво і багатoverшинні. Наявність двох і більше вершин свідчить про неоднорідність сукупності, про поєднання в ній груп з різними рівнями ознаки. Розподіли якісно однорідних сукупностей, як правило, одновершинні. Серед одновершинних розподілів є симетричні (скошені), гостро- і плосковершинні.

Різновиди форм розподілу наведені на Рис. 5.1.

У **симетричному** розподілі рівновіддалені від центру значення ознаки мають однакові частоти, в **асиметричному** – вершина розподілу зміщена. Напрямок асиметрії протилежний напрямку зміщення вершини. Якщо вершина зміщені вліво, то це правостороння асиметрія, і навпаки. Асиметрія виникає внаслідок обмеженої варіації в одному напрямку або під впливом домінуючої причини розвитку, яка веде до зміщення центру розподілу.

Найпростішою мірою асиметрії є відхилення між середньою арифметичною і медіаною чи модою. У симетричному розподілі характеристики центру мають однакові значення $\bar{X} = M_e = M_0$; в асиметричному між ними існують певні розбіжності. При правосторонній асиметрії $\bar{X} > M_e > M_0$, при лівосторонній асиметрії, навпаки, $\bar{X} < M_e < M_0$.

Як узагальнюючі характеристики розподілу використовують **моменти**. За допомогою невеликого їх числа можна описати будь-який розподіл. **Момент розподілу** – це середня арифметична k -го ступеня відхилень $X - A$.

Залежно від величини A моменти поділяють на первинні $A = 0$, центральні $A = x$ і умовні $A = const$. Ступінь k визначає порядок моменту.

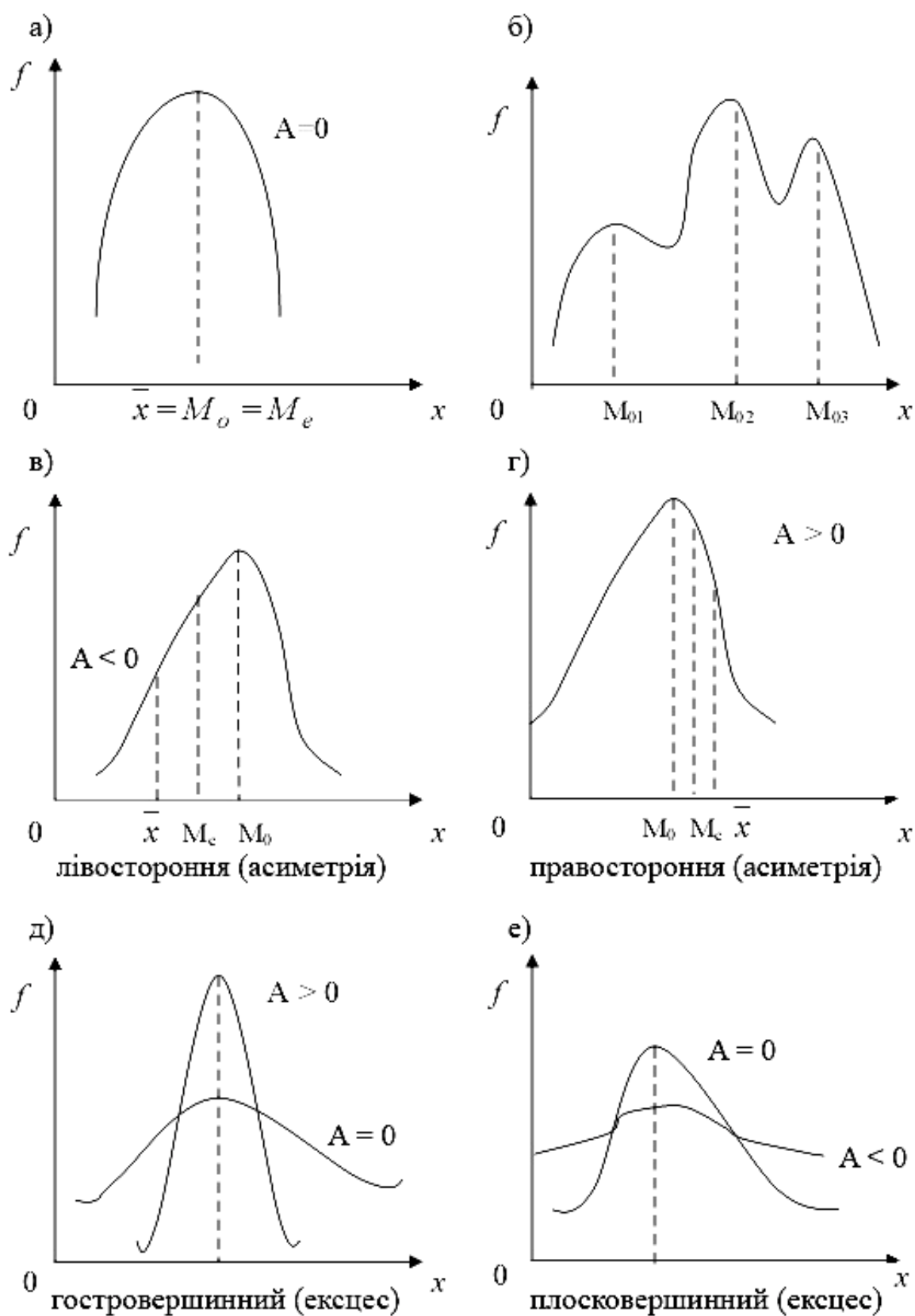


Рис. 5.1 - Різновиди форм розподілу

5. Графічне зображення рядів розподілу

Для графічного зображення рядів розподілу використовують такі графіки як полігон, гістограма, кумулята, огіва, крива концентрації (Лоренца), показникова крива, крива Парето, антимода.

Полігон – графічне зображення варіаційного ряду в прямокутній системі координат, при котрому величина ознаки відкладається на осі абсцис, частоти або частки (щільність розподілу) – на осі ординат.

Частіше полігон застосовують для зображення дискретного варіаційного ряду, але може використовуватись і для інтервального ряду. Графік полігону має такий вигляд:

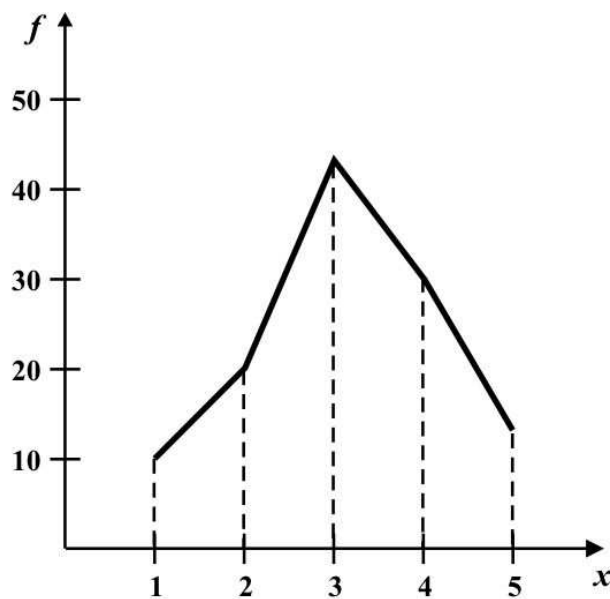


Рис 5.2 – Полігон

Гістограма – графічне зображення інтервального варіаційного ряду у вигляді прямокутників різної висоти, основа яких – відрізки осі абсцис, котрі відповідають інтервалам зміни ознаки. Висоти прямокутників пропорціональні при рівності інтервалів частотам або часткам інтервалів, а при нерівності – щільностям (абсолютним чи відносним). Графік гістограми має такий вигляд:

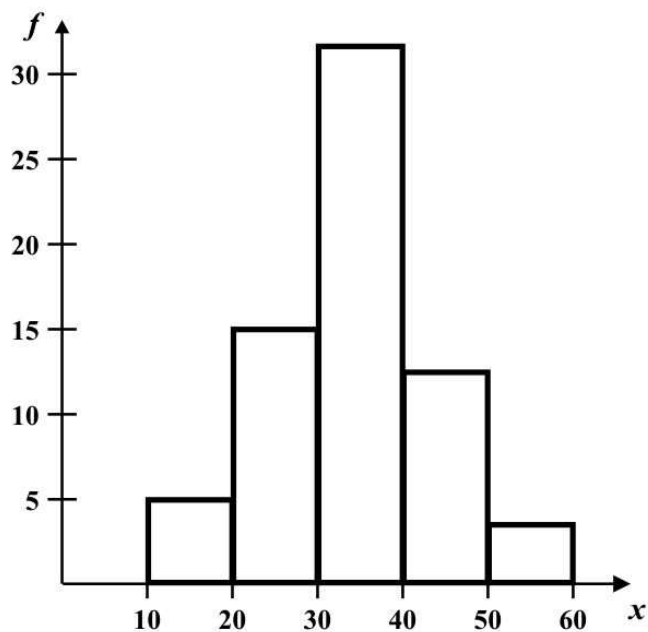


Рис. 5.3 – Гістограма

Кумулятивна крива (кумулята) – графічне зображення варіаційного ряду, складене за нагромадженими частотами або частками.

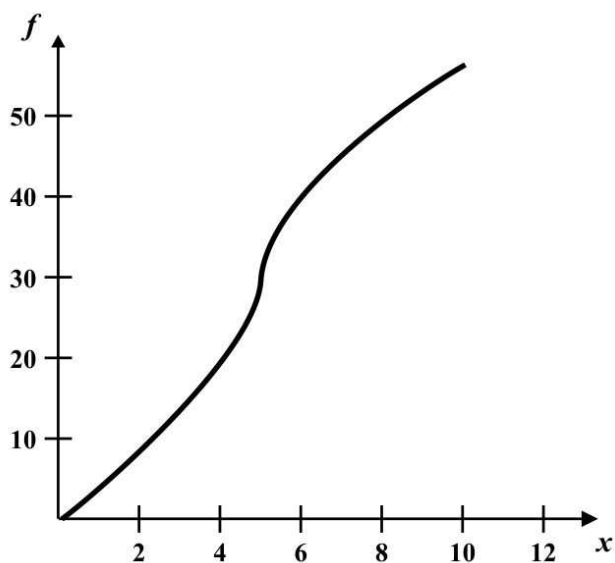


Рис. 5.4 – Кумулята

Огіва – графічне зображення варіаційного ряду, складене аналогічно кумуляті, але на осі ординат наносяться значення ознаки, а на осі абсцис – нагромаджені частоти.

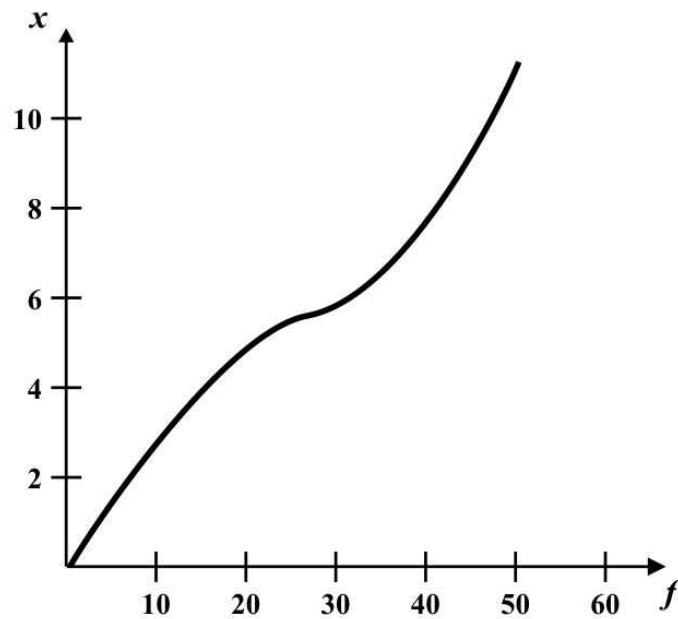


Рис. 5.5 – Огіва

Крива концентрації Лоренца – використовується як показник ступеня оцінки концентрації розподілу частот. Вона має такий вигляд:

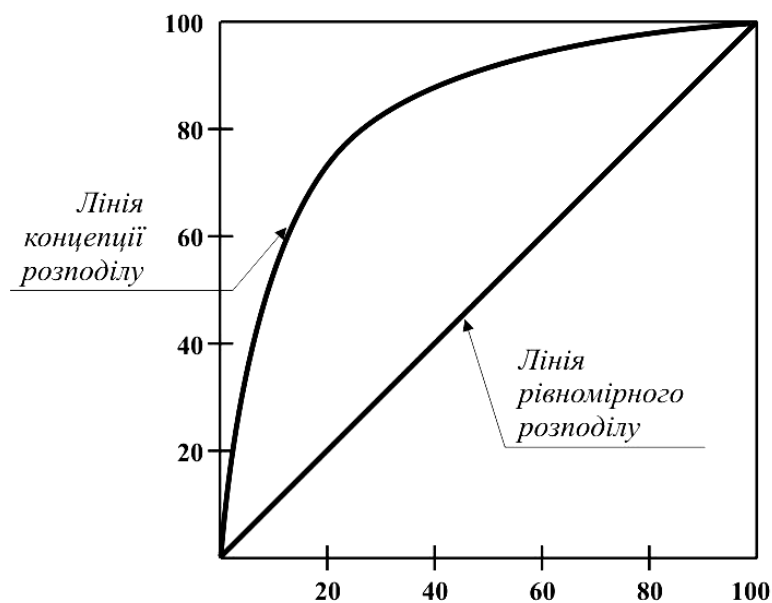


Рис. 5.6 – Крива концентрації Лоренца

До інших видів графічного зображення рядів розподілу відносять: показникову криву, криву Парето і антимоду. Ці графіки мають вигляд:

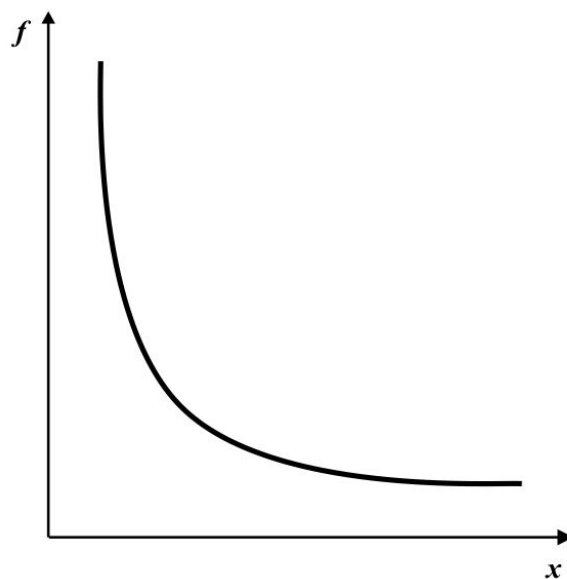


Рис. 5.7 - Показникова крива

Прикладом показникового розподілу є термін служби предметів, які вибувають через аварійність – довговічність посуду в підприємствах громадського харчування, тривалість телефонних викликів, час між простоями верстатів тощо.

Крива Парето показує розподіл доходів у відповідності з їх величиною.

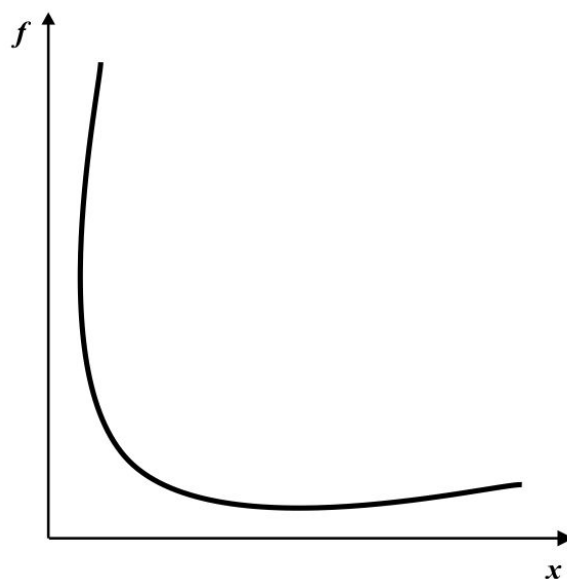


Рис. 5.8 - Крива Парето

Антимода – абсциса нижньої точки, розміщеної в центральній частині U -подібного розподілу ознаки. По обидві сторони антимоди частоти зростають.

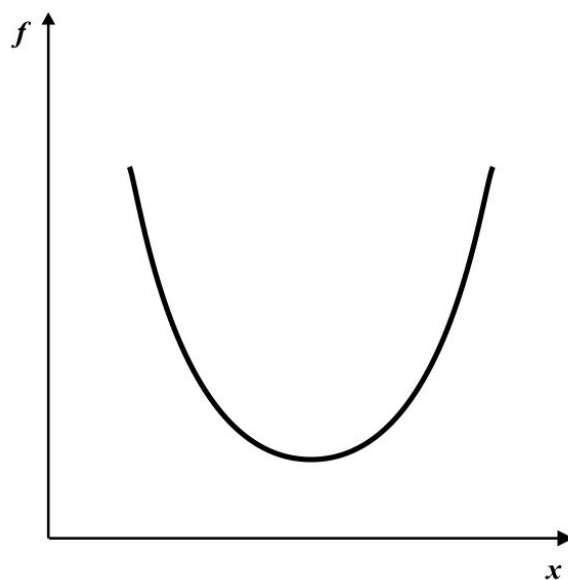


Рис. 5.9 - Антимода



Питання для самоконтролю по темі 5

1. *Дайте визначення терміну «ознака».*
2. *Дайте визначення терміну «варіація».*
3. *Поняття про варіацію ознак.*
4. *Поняття статистичних показників варіації.*
5. *Абсолютні показники варіації.*
6. *Відносні показники варіації.*
7. *Властивості дисперсії.*
8. *Спрощені способи обчислення дисперсії.*
9. *Правило додавання дисперсій.*
10. *Коефіцієнт детермінації.*
11. *Емпіричне кореляційне відношення.*
12. *Поняття про ряди розподілу.*
13. *Види рядів розподілу.*
14. *Форми рядів розподілу.*
15. *Щільність розподілу та інтерполяція.*
16. *Характеристика форм розподілу.*
17. *Криві розподілу та їх типи.*
18. *Графічне зображення рядів розподілу.*

ТЕМА 6. ВИБІРКОВИЙ МЕТОД

Питання для теоретичної підготовки

1. Поняття про вибіркове дослідження та його основні завдання.
2. Вибіркове спостереження, репрезентативність, властивості сукупності.
3. Методи і способи відбору одиниць у вибірккову сукупність.
4. Визначення середньої і граничної помилок та необхідної чисельності вибірки.
5. Мала вибірка.
6. Способи поширення даних вибіркового спостереження на генеральну сукупність.

1. Поняття про вибіркове дослідження та його основні завдання



Спостереження – загальнонауковий спеціальний метод збору первинної інформації шляхом безпосередньої реєстрації подій, явищ і процесів, що відбуваються в певних умовах.

Несуцільне спостереження – тип спостереження, при якому обстеженню підлягають не всі одиниці досліджуваної сукупності, а тільки їх певна частина, достатня для одержання узагальнюючих статистичних характеристик.



Вибіркове спостереження як один з видів несуцільного спостереження є обстеженням певної частини одиниць генеральної сукупності, обов'язково відібраних у випадковому порядку.



Генеральна сукупність (N) – вся підлягаюча вивченню статистична сукупність однорідних об'єктів або явищ суспільного життя, що мають загальні якісні й змінні кількісні ознаки.

У практиці статистичної роботи при вивченні якісної або кількісної ознаки великого обсягу однорідних об'єктів не завжди є можливість обстежити кожний з цих об'єктів. Приклади: якість електроприладів, що поставляються; обстеження сімейних бюджетів; обстеження будівельних матеріалів і конструкцій відповідно до ДСТУ тощо. У цьому випадку досліджуються точні властивості невеликої частини всієї сукупності, відібраної випадковим чином.



Вибіркове спостереження засноване на застосуванні **вибіркового методу**, тобто методу статистичного дослідження, *за допомогою якого обчислюються узагальнюючі показники досліджуваної генеральної сукупності за даними обстеження тільки деякої її частини – вибіркової сукупності (вибірки) (n) на основі положень випадкового відбору.*

Математично теорія вибіркового методу встановлює *залежність* точності результатів вибірки від її виду, застосованого в обстеженні, від числа одиниць, що підлягають спостереженню, і ступеню розсіювання досліджуваних ознак.

Вибіркове спостереження застосовують також у поєднанні із суцільним для поглиблення дослідження, або уточнення і контролю результатів суцільного спостереження.

Вибіркове спостереження проводиться для вирішення наступних **основних завдань**:

- 1) визначення середнього розміру досліджуваної ознаки;
- 2) визначення питомої ваги (частки);
- 3) визначення середньої і граничної помилки вибірки;
- 4) знаходження меж для середньої і частки при повторному і безповторному відборі;
- 5) визначення потрібної чисельності вибірки;
- 6) поширення даних вибіркового спостереження на всю сукупність.

Вибіркове спостереження складається з таких етапів:

- 1) постановка мети спостереження;
- 2) складання програми спостереження і розробка відповідних даних;
- 3) вирішення організаційних питань проведення спостереження;
- 4) визначення відсотка і способу відбору одиниць;
- 5) проведення відбору;
- 6) реєстрація відповідних ознак у відібраних для дослідження одиниць;
- 7) узагальнення даних спостереження та розрахунок їх вибірових характеристик;
- 8) знаходження помилок вибірки;
- 9) перерахунок характеристик вибіркового спостереження на генеральну сукупність.

2. Вибіркове спостереження, репрезентативність, властивості сукупності

Призначення вибіркового спостереження полягає в тому, що при мінімальній кількості обстежених одиниць скорочується час проведення дослідження при досить надійних результатах. Підвищується оперативність статистичної інформації, знижуються помилки реєстрації, з'являється можливість для перевірки даних суцільного обліку.

У порівнянні з іншими методами, що застосовують несучільне спостереження, вибіровий метод має важливі особливості: в основі відбору одиниць для обстеження покладені принципи рівних можливостей залучення у вибірку кожної одиниці N . Саме в результаті дотримання цих принципів виключається утворення вибіркової сукупності тільки за рахунок кращих або гірших зразків.

Вибірка буде правильно відображати властивості всієї сукупності, якщо вибір проводиться випадково, тобто так, що кожна з можливих вибірок заданого обсягу n із сукупності обсягу N має однакову ймовірність бути фактично обраною.

Число таких вибірок (C_N^n) дорівнює:

$$C_N^n = \frac{N!}{n!(N-n)!}, \quad (6.1)$$

Щоб висновок, отриманий шляхом вивчення вибірки, можна було поширити на усю генеральну сукупність, вибірка повинна мати властивості репрезентативності.



Репрезентативність (від франц. *représentatif* – показовий, характерний) у статистиці – головна властивість вибіркової сукупності, що полягає в близькості її характеристик (складу, середніх величин тощо) до відповідних характеристик генеральної сукупності, з якої відібрана (з дотриманням певних правил) вибірка.

Висновки про ступінь репрезентативності складаються на підставі розгляду вибіркової сукупності по двох напрямках:

- По-перше, вона порівнюється з генеральною сукупністю відносно усіх ознак, зафіксованих як в тій, так і в іншій.
- По-друге, висновок про ступінь репрезентативності може бути складений на підставі коливання досліджуваних характеристик у вибірковій сукупності.

Як правило, при вибіркового методі обстеженню підлягає порівняно невелика частина всієї досліджуваної сукупності (звичайно до 5-10%, рідше – 15-25%).

Узагальнюючими показниками генеральної та вибіркової сукупності є:

- у **генеральній сукупності** N частка одиниць, що володіють досліджуваною ознакою називають **генеральною часткою** (p). Середня величина ознаки, що варіює, – **генеральна середня** (\bar{X}), **генеральна дисперсія** (σ^2).
- у **вибірковій сукупності** n частку досліджуваної ознаки називають **вибірковою часткою** (w), середню величину у вибірці – **вибірковою середньою** (\bar{X}), **вибірковою дисперсією** (σ_a^2).



Основне завдання вибіркового дослідження полягає у тому, щоб на підставі характеристик вибіркової сукупності (w або \bar{X}) отримати достовірні судження про показники частки p або \bar{X} у генеральній сукупності.

Властивості сукупності, досліджувані вибіркоким методом, можуть бути *якісними й кількісними*. У першому випадку завдання вибіркового дослідження полягає у визначенні кількості M об'єктів сукупності, що володіють будь-якою ознакою (наприклад, при статистичному контролі часто цікавляться кількістю M дефектних виробів у партії обсягу N). Оцінкою для M постає відношення m/n , де m – число об'єктів з даною ознакою у вибірці обсягу n . У випадку вибіркового дослідження кількісної ознаки мають справу з визначенням середнього значення сукупності.

Відповідь на питання про те, яка за розміром різниця між генеральними і вибілковими узагальнюючими показниками, з якою ймовірністю можна судити про цю різницю, дає *теорія вибіркового методу* на підставі закону великих чисел.

За допомогою цього закону розв'язують два взаємозв'язаних завдання:

- 1) розраховують із заданою ймовірністю межі можливих відхилень вибіркового від відповідного показника у генеральній сукупності;
- 2) визначають ймовірність того, що розмір можливих відхилень вибіркового показника від генерального не перевищить встановленої межі.

При масовому спостереженні, розподіл емпіричних частот більшості явищ підпорядковується закону нормального розподілу.

Доведено, що за нормальним розподілом більша частина величин зосереджена навколо генеральної середньої. Близько 68,3 % чисельності вибіркової середньої буде знаходитись у межах $\pm \sigma$ генеральної середньої; 95,4 % цієї чисельності знаходиться у межах $\pm 2\sigma$ і 99,7 % не вийде за межі $\pm 3\sigma$.

Нормальний розподіл показує частоту виникнення помилок даного розміру середньої.

3. Методи і способи відбору одиниць у вибірку сукупність



Способом відбору називається система організації відбору одиниць з генеральної сукупності. Розрізняють два методи відбору одиниць у вибірку сукупність: повторний і безповторний.

Повторним називається такий метод відбору, при якому кожна раніше відібрана одиниця повертається у генеральну сукупність і може знову брати участь у вибірці.

Безповторним називається такий метод відбору, при якому кожна раніше відібрана одиниця не повертається у генеральну сукупність і у подальшій вибірці участі не бере.

Оскільки безповторний відбір охоплює постійно нові одиниці сукупності, а повторний – одну й ту ж сукупність, тому безповторний відбір дає більш точні результати.

Повторний і безповторний методи відбору залежно від характеру одиниці відбору застосовуються у поєднанні з іншими видами відбору.

У практиці статистичного дослідження використовуються **три види відбору**:

- 1) індивідуальний – відбір окремих одиниць сукупності;
- 2) груповий (серійний) – відбір груп (серій) одиниць;
- 3) комбінований – комбінація індивідуального і групового.

За **способом відбору** одиниць для обстеження розрізняють такі **види вибіркового спостереження**:

- 1) власне випадкова вибірка;
- 2) механічна вибірка;
- 3) типова (районована) вибірка;
- 4) серійна (гніздова) вибірка;
- 5) комбінована вибірка;
- 6) проступінчаста й багатоступінчаста вибірка;
- 7) однофазна й багатозфазна вибірка;
- 8) інші види вибірки.

Власне випадковою називається така вибірка, при якій відбір одиниць з генеральної сукупності є випадковим. Часто для цього застосовують жеребкування або таблицю випадкових чисел.

Механічна вибірка – це послідовний відбір одиниць через рівні проміжки в порядку визначеного розположення їх в генеральній сукуп-

ності, або в якому-небудь переліку. Інтервали відбору визначаються у відповідності з часткою відбору одиниць (кожна п'ята, десята, сота тощо).

При типовому відборі генеральну сукупність поділяють на однорідні групи за певною ознакою, райони, зони. Потім з кожної групи випадковим або механічним способом відбирають певну кількість одиниць, пропорційно частці групи в загальній сукупності.

При серійній (гніздовій) вибірці відбір одиниць проводять цілими групами (серіями, гніздами) сукупності в межах яких обстежують всі одиниці без винятку. Серії для спостереження відбирають випадково, частіше неповторним способом механічної вибірки.

Комбінованою називається така вибірка, коли комбінують два або кілька видів вибірок. Перш за все, комбінують суцільне і вибіркове спостереження. В даному випадку, за основною програмою обстежується генеральна сукупність, а за додатковою – вибіркова.

Одноступінчастою називається вибірка, коли із досліджуваної сукупності зразу відбираються одиниці або серії одиниць для безпосереднього обстеження.

Багатоступінчаста вибірка передбачає поступове вилучення з генеральної сукупності спочатку укрупнення груп одиниць, потім груп менших за обсягом, і так до тих пір, доки не відберуть відповідні групи або одиниці, які будуть досліджуватись. Вибірка може бути двох-, трьох і більше ступінчастою.

Якщо необхідні дані можна отримати на основі вивчення всіх первинно відібраних одиниць, застосовують **однофазну вибірку**, а якщо тільки на основі деякої її частини, відібраної так, що вона складає підвибірку із початково роведеної вибірки – **багатофазну**.

Багатофазною називається така вибірка, коли одні відомості збираються від всіх одиниць відбору, потім відбираються ще деякі одиниці

і обстежуються за більш широкою програмою. При багатофазній вибірці на кожній фазі зберігається одна і та ж одиниця відбору.

Розрахунок помилок репрезентативності багатоступінчастої і багатофазної вибірок проводиться для кожної ступені і фази окремо.

Бувають випадки, коли необхідно застосувати інші види відбору, такі як взаємопроникаючі і квантильні вибірки, направлений відбір, моментні спостереження, або скористатись малою вибіркою.

Взаємопроникаючою називається така вибірка, коли із однієї генеральної сукупності проводять одним і тим же способом декілька незалежних вибірок.

Взаємопроникаючі вибірки завжди проводять різні, незалежні один від одного дослідники, що дозволяє порівнювати підсумки по всіх частинах і забезпечити взаємну перевірку їх роботи. Взаємопроникаючі вибірки дають незалежні одна від одної оцінки значень досліджуваної сукупності, і якщо результати різних вибірок близькі між собою, то такі оцінки дуже переконливі.

Помилки взаємопроникаючих вибірок визначаються за формулами типової пропорційної вибірки.

Квантильні вибірки застосовують тоді, коли виникає потреба дослідження даних суцільного спостереження за додатковою програмою.

Для проведення квантильної вибірки рангують потрібну варіаційну ознаку і за її нагромадженими частотами будують огіву. За огіву механічним способом відбирають потрібну частину одиниць для дослідження цієї ж ознаки. Якщо огіва вибіркової сукупності добре відтворює огіву генеральної сукупності, то помилка репрезентативності буде мінімальною.

Направлений відбір використовують тоді, коли за відомим середнім значенням ознаки в генеральній сукупності вибіркова сукупність повинна характеризувати її структуру за іншими ознаками.

Направлений відбір передбачає проведення відбору таким чином, щоб середній розмір відібраних одиниць дорівнював середньому розміру одиниць всієї сукупності. В тому випадку, коли заміна однієї одиниці іншою призводить до наближеної рівності середніх генеральної і вибіркової сукупностей, вибірку вважають врівноваженою і репрезентативною за всіма іншими ознаками сукупності. Таким чином, **направленим відбором** називається врівноваження за однією ознакою для вибіркового дослідження інших ознак.

Помилку вибірки направленої відбору визначають в залежності від способу проведення відбору одиниць до врівноваження.

Моментне спостереження використовується для вивчення використання робочого часу робітниками або часу роботи устаткування.

4. Визначення середньої і граничної помилок та необхідної чисельності для різних видів вибірок

Для вибіркового спостереження властиві помилки реєстрації і помилки репрезентативності.

Помилки репрезентативності становлять різницю між середніми і відносними показниками вибіркової сукупності та відповідними показниками генеральної сукупності. Вони поділяються на систематичні та випадкові.

Систематичні помилки репрезентативності зумовлені порушенням принципів проведення вибіркового спостереження.

Випадкові помилки репрезентативності зумовлені тим, що вибірка сукупність не відображає точно середні і відносні показники генеральної сукупності.

Визначення величини випадкових помилок репрезентативності є одним з головних завдань теорії вибіркового методу.

Для узагальнюючої характеристики помилки вибірки вираховують **середню помилку репрезентативності** (μ) і називають ще стандартом.

Для визначення **середньої помилки репрезентативності** власне випадкової і механічної вибірки застосовують чотири формули для повторного і без повторного відбору (табл. 6.1).

Таблиця 6.1 – Середня помилка репрезентативності

Спосіб відбору	При визначенні середньої	При визначенні частки
Повторний	$\mu = \sqrt{\frac{\sigma^2}{n}}$	$\mu = \sqrt{\frac{w(1-w)}{n}}$
Безповторний	$\mu = \sqrt{\frac{\sigma^2}{n} \left(1 - \frac{n}{N}\right)}$	$\mu = \sqrt{\frac{w(1-w)}{n} \left(1 - \frac{n}{N}\right)}$

де μ – середня помилка репрезентативності;
 σ^2 – середній квадрат відхилень у вибірці;
 n – чисельність вибіркової сукупності;
 N – чисельність генеральної сукупності;
 n / N – частка обстеженої частини вибіркової сукупності;
 $1 - n / N$ – необстежена частка генеральної сукупності;
 w – частка даної ознаки у вибірці;
 $(1 - w)$ – частка протилежної ознаки у вибірці.

На практиці частіше використовують безповторний відбір, який гарантує більшу точність результатів.

Для узагальнюючої характеристики помилки вибірки поряд із середньою розраховують ще **граничну помилку вибірки**.

При вибіркового спостереженні розмір граничної помилки репрезентативності Δ може бути більшим, дорівнювати або бути меншим

від середньої помилки репрезентативності μ . Тому величину граничної помилки репрезентативності обчислюють з певною ймовірністю P , якій відповідає t -разове значення μ . Із введенням показника кратності помилки t , формула граничної помилки репрезентативності матиме вигляд:

$$\Delta = t * \mu; \quad t = \frac{\Delta}{\mu}, \quad (6.2)$$

де μ – середня помилка вибірки;

t – коефіцієнт довір'я, який залежить від ймовірності визначення граничної помилки.

Ймовірність відхилень вибіркової середньої від генеральної середньої при достатньо великому обсязі вибірки і обмеженій дисперсії генеральної сукупності підпорядковується закону нормального розподілу. Ймовірність цих відхилень при різних значеннях t визначається за спеціальною формулою.

Значення цих відхилень при різних значеннях t табульовані і приводяться у спеціальних таблицях, наприклад:

для $t = 1$ $p(\Delta \leq \mu) = 0,683$; для $t = 2$ $p(\Delta \leq \mu) = 0,954$;

для $t = 3$ $p(\Delta \leq \mu) = 0,997$; для $t = 4$ $p(\Delta \leq \mu) = 0,999$.

Гранична помилка вибірки дає можливість встановити, в яких межах знаходиться величина генеральної середньої або частки.

З теореми Чебишева знаходять, що:

$$\bar{x} - \tilde{x} = \pm \Delta x; \quad \tilde{x} - \Delta x \leq \tilde{x} \leq \tilde{x} + \Delta x, \quad (6.3)$$

Додаючи граничну помилку вибірки до вибіркової частки і віднімаючи граничну помилку вибірки від вибіркової частки, знаходять межі генеральної частки:

$$p - w = \pm \Delta p ; \quad w - \Delta p \leq p \leq w + \Delta p , \quad (6.4)$$

На основі формул граничної помилки вибірки розв'язують наступні **завдання**:

- 1) визначають довірчі межі генеральної середньої і частки з прийнятою ймовірністю;
- 2) визначають ймовірність того, що відхилення між вибірковими і генеральними характеристиками не перевищать визначену величину;
- 3) визначають необхідну чисельність вибірки, яка із заданою ймовірністю забезпечить прийняту точність вибірових показників.

При організації проведення вибіркового спостереження важливе значення має правильне визначення необхідної чисельності вибірки, яка з відповідною ймовірністю забезпечить встановлену точність результатів спостереження.

Чисельність вибірки залежить від наступних **чинників**:

- 1) від варіації досліджуваної ознаки. Чим більша варіація, тим більшою повинна бути чисельність вибірки, і навпаки;
- 2) від розміру можливої граничної помилки вибірки. Чим менший розмір можливої помилки, тим більшою повинна бути чисельність вибірки. Існує правило, якщо помилку потрібно зменшити в три рази, то чисельність вибірки збільшують у дев'ять разів;
- 3) від розміру імовірності, з якою гарантуватимуть результати вибірки. Чим більша імовірність, тим більшою повинна бути чисельність вибірки;
- 4) від способу відбору одиниць у вибіркову сукупність для обстеження.

Основні формули для знаходження необхідної чисельності вибірки для власне випадкової і механічної вибірки наведені у табл. 6.2.

Таблиця 6.2 – Необхідна чисельність вибірки

Способи відбору	Чисельність вибірки	
	При визначенні середньої	При визначенні частки
Повторний	$n = \frac{t^2 \sigma_X^2}{\Delta_x^2}$	$n = \frac{t^2 w(1-w)}{\Delta_w^2}$
Безповторний	$n = \frac{N t^2 \sigma^2}{N \Delta_X^2 + t^2 \sigma^2}$	$n_w = \frac{N t^2 w(1-w)}{N \Delta_w^2 + t^2 w(1-w)}$

Цей розрахунок підтверджує, що при інших рівних умовах, обсяг вибірки при безповторному відборі завжди буде менший, ніж при повторному.

5. Мала вибірка

Мала вибірка – несуцільне статистичне обстеження, при якому вибіркова сукупність утворюється зі порівняно невеликої кількості одиниць сукупності. Малою вибіркою називається вибіркова сукупність, яка складається з порівняно невеликої кількості одиниць (20 - 30). На практиці іноді доводиться обмежуватись малою кількістю спостережень (при перевірці якості продукції, зв'язаної із знищенням продукції, яку перевіряють). Математичною статистикою доведено, що і при малих вибірках характеристики вибіркової сукупності можна поширити на генеральну.

Англійський вчений Стюдент винайшов закон розподілу відхилень вибірових середніх від генеральної середньої для малих вибірок.

Опираючись на цей закон, він склав спеціальні таблиці, в яких наводяться значення критерію t для малих вибірок.

При **малих вибірках** розподіл вибірових середніх і помилок вибірки відрізняється від нормального. Тому для оцінки результатів малої вибірки використовують дещо видозмінені формули. Середня помилка малої вибірки розраховується за формулою:

$$\mu_{\hat{a}} \approx \sqrt{\frac{\sigma_{\hat{a}}^2}{n}}, \quad (6.5)$$

де $\sigma_{\hat{a}}^2$ – дисперсія малої вибірки.

Гранична помилка малої вибірки:

$$\Delta_{\hat{a}} = t * \mu_{\hat{a}}, \quad (6.6)$$

При цьому значення коефіцієнта довіри t залежить не тільки від заданої довірчої ймовірності, але й від чисельності одиниць вибірки n .

Для окремих значень t і n довірна ймовірність малої вибірки визначається по спеціальних таблицях **Стюдента**, які приводяться в підручниках по «Математичній статистиці».

6. Способи поширення даних вибіркового спостереження на генеральну сукупність

Кінцевою практичною метою всякого вибіркового спостереження є поширення його характеристик на генеральну сукупність.

Існують два способи розповсюдження даних вибіркового спостереження:

- 1) спосіб прямого перерахування;
- 2) спосіб коефіцієнтів.

Спосіб прямого перерахування застосовують в тому випадку, коли на основі вибірки розраховують об'ємні показники генеральної сукупності, використовуючи для цього вибіркові середню або частку. В першому випадку середній розмір ознаки, визначений в результаті вибіркового спостереження, множиться на кількість одиниць генеральної сукупності.

Встановивши, наприклад, в результаті вибіркового спостереження продуктивності праці в промисловості, що середній виробіток на одного працюючого в області становить 2 500 грн./міс., і знаючи, що усього в області у цій галузі промисловості 20 000 працюючих, можемо отримати величину валового продукту: $2\,500 * 20\,000 = 50\,000\,000$ грн. або 50 млн. грн. Якщо при цьому відомо, що середня помилка вибірки з певною ймовірністю дорівнює ± 20 грн./міс., тож генеральна середня з тією ж ймовірністю коливається в межах від 2 480 до 2 520 грн./міс., а загальний валовий продукт цієї галузі промисловості з урахуванням помилки вибірки буде коливатися від 49,6 млн. грн. до 50,4 млн. грн.

Спосіб поправочних коефіцієнтів застосовується в тих випадках, коли вибіркове спостереження проводиться з метою перевірки і уточнення результатів суцільного спостереження. В даному випадку, співставляючи дані вибіркового спостереження із суцільним, вираховують поправочний коефіцієнт, який використовують для внесення поправок в матеріали суцільних спостережень.

Приклад. По закінченні перепису мобільних телефонів оператор мобільного зв'язку проводять 10 %-ве контрольне вибіркове обстеження апаратів, які знаходяться в особистій власності населення, і якщо в результаті контрольного обходу виявляється недооблік, тоді дані перепису коригуються на відсоток недообліку.

Припустимо, що в результаті пристроїв, що знаходяться в особистій власності населення району, нараховано 15 000 одиниць, в тому числі в населених пунктах, де проводиться 10 %-ва вибірка – 1 200 одиниць.

В результаті контрольних обходів в цих же населених пунктах було обчислено 1 215 одиниць. Тобто, при переписі було недообліковано 15 пристроїв, що складає 1,25 % (15 від 1200). За допомогою цього поправочного коефіцієнта уточнюють матеріали перепису суцільного спостереження.

В даному випадку по району всього недообліковано 187 мобільних пристроїв ($\frac{1,25 \cdot 15000}{100}$). Отже загальна кількість пристроїв, що знаходяться в особистій власності населення району, з поправкою на недооблік складає 1 387 (1200 + 187).



Питання для самоконтролю по темі 6

1. Що таке вибіркове спостереження?
2. Умови застосування вибіркового спостереження.
3. Дайте визначення терміну генеральна сукупність.
4. Дайте визначення терміну вибірка сукупність.
5. Узагальнюючі характеристики генеральної та вибіркової сукупності.
6. Способи відбору одиниць із генеральної сукупності.
7. Повторний і безповторний відбір.
8. Випадкова, механічна, типова, серійна вибірки.
9. Особливості малої вибірки.
10. Види помилок при вибіркового спостереженні.
11. Визначення помилок вибірки для середньої і частки.
12. Визначення необхідної чисельності вибірки.
13. Поширення даних вибіркового дослідження на генеральну сукупність.

ЗМІСТОВИЙ МОДУЛЬ 1.2. МЕТОДИ СТАТИСТИКИ

ТЕМА 7. СТАТИСТИЧНА ПЕРЕВІРКА ГІПОТЕЗ

Питання для теоретичної підготовки

1. Загальні поняття, вибір типу критичної ділянки.
2. Перевірка гіпотези про належність спостережень, які видрізняються від досліджуваної генеральної сукупності.
3. Перевірка гіпотези про відповідність емпіричного розподілу типу кривих нормального розподілу.
4. Перевірка гіпотези про величину центру розподілу.

1. Загальні поняття, вибір типу критичної ділянки

Під **статистичною гіпотезою** розуміють різні передбачення щодо виду або параметрів розподілу випадкової перемінної, які можна перевірити, спираючись на результати спостережень у власне випадковій вибірці. Статистична перевірка гіпотез носить ймовірний характер, так як їх висновки ґрунтуються на вивченні властивостей розподілу випадкової перемінної за даними вибірки, а тому завжди є ризик допустити помилку. Однак, з допомогою статистичної перевірки гіпотез можна визначити ймовірність прийняття помилкового висновку. Якщо його ймовірність незначна, то вважається, що застосований критерій забезпечує малий ризик помилки.

При перевірці гіпотез існує можливість допустити помилку двоякого роду:

- а) гіпотеза, яка перевіряється (нульова гіпотеза « H_0 »), в дійсності є вірною, але результати перевірки приводять до відмови від неї;
- б) гіпотеза, яка перевіряється, в дійсності є помилковою, але результати перевірки проводять до її прийняття.

Для побудови статистичного критерію, який дозволяє перевірити певну гіпотезу, потрібно виконати наступне:

1. Сформулювати гіпотезу « H_0 », яка перевіряється. Поряд з нею висувається альтернативна гіпотеза;
2. Вибрати рівень значущості « α », який контролюватиме допустиму ймовірність помилки першого роду;
3. Визначити критичну ділянку допустимих значень;
4. Прийняти відповідне рішення.

Рівнем значущості називається таке мале значення ймовірності попадання критерію в критичну ділянку при умові справедливості гіпотези, що появу цієї події можна розцінювати як наслідок суттєвого розходження висунутої гіпотези з результатами вибірки. Як правило, рівень значущості приймають рівним 0,01 або 0,05. Виходячи з величини рівня значущості визначають критичну ділянку, під якою розуміють таку ділянку значень вибіркової характеристики, попадаючи в котру вони будуть свідчити про те, що гіпотеза яка перевірялась повинна бути відкинута.

До **критичної ділянки** відносяться такі значення, поява яких за умови правильності гіпотези була б малоімовірною. Припустимо, що обчислене за емпіричними даними значення критерію попало в критичну ділянку, тоді за умови ймовірності перевіркової гіпотези « H_0 » ймовірність цієї події буде не більша значущості « α ». Оскільки « α » вибирається достатньо малим, то така подія є малоімовірною, і, отже, перевірна гіпотеза « H_0 » може бути відкинута. Якщо ж досліджуване значення критерію не належить до критичної ділянки, а знаходиться в ділянці допустимих значень, то перевірна гіпотеза « H_0 » не відкидається.

Ймовірність попадання критерію в ділянку допустимих значень при справедливості перевіркової гіпотези « H_0 » дорівнює $1 - \alpha$. Потрібно враховувати, що попадання критерію в ділянку допустимих значень не означає строгого доведення гіпотези « H_0 ». Воно лише вказує, що між висунутою гіпотезою і результатами вибірки не існує суттєвих розбіжностей. Із зменшенням рівня значущості, зменшується ймовірність забра-

кувати перевіро́чну гіпотезу, коли вона вірна, тобто менша ймовірність допустити помилку першого роду. Але при цьому розширюється ділянка допустимих значень і, значить, збільшується ймовірність зробити помилки другого роду.

За даним рівнем значущості можливо по-різному встановлювати критичну область, яка гарантує цей рівень. Нехай ми задаємося рівнем значущості $\alpha = 0,05$.

В якості критичної ділянки, яка відповідає даному рівню значущості, можемо взяти (для нормального розподіленого критерію):

1. Ділянку великих позитивних відхилень так, щоб

$$P_I = P(x > \bar{x} + t_I \mu) = 0,05, \quad (7.1)$$

Тоді
$$P_I = \frac{1}{\sqrt{2\pi}} * \int_t^{\infty} e^{-\frac{t^2}{2}} dt = \frac{1}{2} - \frac{1}{\sqrt{2\pi}} * \int_0^t e^{-\frac{t^2}{2}} dt, \quad (7.2)$$

звідки
$$\frac{1}{\sqrt{2\pi}} * \int_0^t e^{-\frac{t^2}{2}} dt = \frac{1}{2} - 0,05 = 0,45$$

і за таблицею нормованої функції Лапласа знаходимо, що $t_I = 1,64$.

2. Ділянка великих від'ємних відхилень

$$P_{II} = P(x < \bar{x} - t_I \mu) = 0,05, \text{ тобто } P_{II} = P(x < \bar{x} - 1,64\mu), \quad (7.3)$$

3. Ділянка великих за абсолютною величиною відхилень

$$P_{III} = P(|x - \bar{x}| > t_{III} \mu) = 0,05, \text{ тобто } 0,05 = 1 - 2 \frac{1}{\sqrt{2\pi}} * \int_0^t e^{-\frac{t^2}{2}} dt, \quad (7.4)$$

Тоді
$$\frac{1}{\sqrt{2\pi}} * \int_0^t e^{-\frac{t^2}{2}} dt = \frac{1 - 0,05}{2} = 0,45.$$

За таблицею нормованої функції Лапласа знаходимо, що $t_{III} = 1,96$.

4. Ділянка малих за абсолютною величиною відхилень

$$P_{IV} = P(|x - \bar{x}| < t_{IV} \mu) = 2 \frac{1}{\sqrt{2\pi}} \int_0^t e^{-\frac{t^2}{2}} dt = 0,05, \quad (7.5)$$

$$\text{Тоді} \quad \frac{1}{\sqrt{2\pi}} * \int_0^t e^{-\frac{t^2}{2}} dt = 0,025.$$

Величина t , відповідна ймовірність 0,025, дорівнює 0,65.

2. Перевірка гіпотези про належність спостережень, які видрізняються від досліджуваної генеральної сукупності

У складі зібраних даних можуть зустрічатись поодинокі спостереження, в яких зареєстровані значення ознаки які помітно відрізняються від загального рівня. Вони виникають в результаті:

- а) помилок спостереження;
- б) випадкового збігу різного роду окремих несуттєвих обставин;
- в) порушення однорідності досліджуваної сукупності.

Суб'єктивне відкидання варіантів значень ознаки, які різко відрізняються від середньої арифметичної, не мають ніякого принципового виправдання. Для виключення спостережень які виділяються із подальшої обробки потрібне застосування обґрунтованих критеріїв. Припустимо, що розподіл результатів спостережень, які проводяться в звичайних умовах, відповідає нормальному закону з параметрами « \bar{x} » і « σ ». УВ результаті проведення однієї із серій спостережень отримані « n » значень $x_1 \leq x_2 \leq x_3 \dots \leq x_n$, серед яких максимальне значення « x_n » (або мінімальне « x_1 ») різко відрізняється за своєю величиною від решти « $n - 1$ » спостережень. Виникає питання, чи відносяться ці значення до даної, яка спостерігається за певних умов, генеральної сукупності, чи вони виникли в результаті

яких-небудь екстраординарних обставин? Нульовою гіпотезою в даному випадку є передбачення про те, що « x_n » та « x_1 » належать до тієї самої сукупності, як і всі інші « $n - 1$ » спостережень, тобто « x_n » або « x_1 » не є результатом помилки спостереження або зміни загальних умов формування рівня ознаки в сукупності.

Перевірка цієї гіпотези заключається в тому, що « x_n » та « x_1 » порівнюються за величиною з деякою критичною межею « x ». Якщо спостереження, яке виділяється, є найбільшим, то « x_n » порівнюється з верхньою допустимою межею, вибраною таким чином, щоб ймовірність, перевершивши її, дорівнювала рівню значущості, тобто в даному випадку маємо справу з критичною ділянкою виду $P(x_n > \bar{x} + t\sigma) = \alpha$. Гіпотеза « H » забраковується, якщо « x_1 » перевищує за величиною вказану межу. Якщо ж спостереження, яке виділяється, є найменшим « x_1 », то його порівнюють з нижньою допустимою межею, яка рівна « $\bar{x} - t\sigma$ », тобто $P(x_1 < \bar{x} - t\sigma) = \alpha$. Якщо ж випробуванню одночасно підлягає і максимальне, й мінімальне значення ознаки, то критична ділянка буде мати вигляд $P(|x - \bar{x}| > t\sigma) = \alpha$.

Приклад.

Таблиця 7.1 – Умовні дані для перевірки гіпотези

Кількість спостережень	Мінімальне значення		Максимальне значення		Різниця суміжних значень		Середнє значення	Середнє квадратичне відхилення у вибірці
	x_1	x_2	x_{n-1}	x_n	$x_2 - x_1$	$x_n - x_{n-1}$	\bar{x}	S
100	10	16	140	188	6	48	67,3	45,3

Використовуємо критичну ділянку виду $P(x_n > \bar{x} + t\sigma)$, так як мінімальне значення « x_1 » мало відрізняється від наступного за ним значення у рангованому ряду, тобто перевіримо, чи належить спостереження, яке виділяється, $x_n = 188$ до сукупності результатів, що розглядаються. При

рівні значущості $\alpha = 0,01$ значення нормованої функції Лапласа для критичної ділянки, яка розглядається, дорівнює:

$$\frac{1}{\sqrt{2\pi}} * \int_0^t e^{-\frac{t^2}{2}} dt = \frac{1}{2} - 0,01 = 0,49.$$

Цьому значенню у таблиці нормованої функції Лапласа відповідає $t = 2,33$. Тоді верхня допустима межа значень ознаки, яка не може бути перевищена з ймовірністю 0,99, дорівнюватиме $67,3 + 2,33 * 45,3 = 172,8$. Значення $x_n = 188$ виходить за розраховану межу, а тому з ймовірністю 0,99 можна вважати, що $x_n = 188$ не належить до досліджуваної сукупності і повинно бути виключене з подальших розрахунків.

Часто трапляються випадки, коли параметри генеральної сукупності « \bar{x} » і « σ » невідомі, і для перевірки гіпотези про спостереження, які виділяються, використовуються параметри отримані при вибірковому спостереженні. Проте належить враховувати, особливо при малих вибірках, що ці оцінки є не цілком надійними. У зв'язку з цим для відкидання спостережень, що виділяються за даними малих вибірок, використовують критерій Ф. Груббса.

Критерій Ф. Груббса базується на відношенні двох сум квадратів відхилень.

1. Для випробування найбільшого спостереження, яке виділяється у вибірці обсягом « n » із нормально розподіленої сукупності, розраховують відношення

$$\frac{\sum_{i=1}^{n-1} (x_i - \bar{x}_n)^2}{\sum_{i=1}^n (x_i - \bar{x})^2}, \quad (7.6)$$

$$\text{де } x_1 \leq x_2 \leq x_3 \dots \leq x_n; \quad \bar{x}_n = \frac{\sum_{i=1}^{n-1} x_i}{n-1}; \quad \bar{x} = \frac{\sum_{i=1}^n x_i}{n}.$$

2. Для випробування найменшого спостереження, яке виділяється у вибірці обсягом « n », розраховується відношення:

$$\frac{\sum_{i=1}^n (x_i - \bar{x}_1)^2}{\sum_{i=2}^n (x_i - \bar{x})^2}, \quad (7.7)$$

де $\bar{x}_1 = \frac{\sum_{i=2}^n x_i}{n-1}$.

Обчислена величина відношення « K_p » порівнюється з табличною величиною « K_T » при певному числі спостережень і завданому рівні значущості. « K_T » характеризує ту граничну величину, яка з ймовірністю $(1 - \alpha)$ пояснюється випадковими причинами. Якщо « K_p » дорівнює або менше табличного, то найменше або найбільше спостереження відкидаються. Якщо « K_p » < « K_T », то ймовірність того, що розходження у сумах квадратів відхилень пояснюється випадковими причинами, дорівнює рівню значущості α і в силу малої ймовірності вважається подією гранично не можливою. В такому випадку спостереження, які виділяються, потрібно відкинути і для подальшого обчислення використати спостереження « $n - 1$ », що залишилися.

Розглянемо застосування критерію Ф. Груббса за наступними Даними: відхилення деталей від нормальної ваги (г): 0,06; 0,09; 0,10; 0,11; 0,13 0,14; 0,15; 0,16; 0,24. Потрібно визначити, чи не містять результати помилки спостереження, передбачаючи, що розподіл ваги деталей в генеральній сукупності відповідає закону нормального розподілу.

Середні відхилення по усіх деталях:

$$\bar{x} = \frac{\sum x}{n} = \frac{1,18}{9} = 0,1311 \text{ г.}$$

Виключаємо максимальне відхилення 0,24:

$$\bar{x}_n = \frac{\sum x - x_n}{n-1} = \frac{1,18 - 0,24}{9-1} = \frac{0,94}{8} = 0,1175 \text{ а}.$$

Сума квадратів відхилень від « \bar{x} »:

$$\sum (x - \bar{x})^2 = 0,021289.$$

Сума квадратів відхилень « $n - 1$ » значень від « \bar{x}_n »:

$$\sum (x - \bar{x}_n)^2 = 0,007952.$$

Відношення двох сум квадратів відхилень:

$$K_p = \frac{\sum (x - x_n)^2}{\sum (x - \bar{x})^2} = \frac{0,007952}{0,021289} = 0,3735.$$

Таблиця 7.2 – Витяг із таблиці Ф. Груббса

Число спостережень (<i>n</i>)	При рівні значущості (<i>α</i>)		Число спостережень (<i>n</i>)	При рівні значущості (<i>α</i>)	
	0,01	0,05		0,01	0,05
3	0,0001	0,0027
4	0,0100	0,0494	15	0,4401	0,5559
5	0,0442	0,1270
6	0,0928	0,2032	20	0,5393	0,6379
7	0,1447	0,2696
8	0,1948	0,3261	25	0,6071	0,6923
9	0,2411	0,3742
10	0,2831	0,4154

При числі спостережень $n = 9$ і рівні значущості 1 % $K_T = 0,2411$, тобто $K_p > K_T$ ($0,3735 > 0,2411$), тому відхилення від номінальної ваги 0,24 г не можна віднести до помилок спостереження.

Одним з етапів статистичної обробки результатів спостережень є перевірка гіпотези про можливість віднесення емпіричного розподілу, отриманого за даними вибірки, до типу кривих нормального розподілу.

3. Перевірка гіпотези про відповідність емпіричного розподілу типу кривих нормального розподілу

Між частотами емпіричного і теоретичного розподілу завжди є певне розходження. В деяких випадках ці розходження не є суттєвими і пояснюються випадковостями вибірки. В інших випадках ці розходження пояснюються тим, що розподіл ознаки в генеральній сукупності не відповідає нормальному розподілу. Для того щоб дати обґрунтовану відповідь про причини розходжень емпіричних і теоретичних частот, використовуються *критерії згоди*.

Одним з найпоширеніших критеріїв згоди є **критерій хи-квадрат К. Пірсона (χ^2)**:

$$\chi^2 = \sum \frac{(f_i - f_i')}{f_i'} , \quad (7.8)$$

де f_i , f_i' – відповідно частоти емпіричного і теоретичного розподілу у певному її інтервалі.

Із збільшенням різниці між фактичними і теоретичними частотами зростає величина критерію К. Пірсона. Щоб відрізнити суттєві значення « χ^2 » від значень, які виникають в результаті випадковостей вибірки, вираховане значення критерію порівнюється з табличним значенням « χ_{α}^2 » при відповідному числі ступенів вільності і завданим рівнем значущості. Рівень значущості обирається таким чином, що $P(\chi_{\alpha}^2 > \chi_{\alpha}^2) = \alpha$ (величина α приймається 0,01 або 0,05).

Визначаючи значення критерію К. Пірсона за даними конкретної вибірки, можна зустрітись з такими варіантами:

1. $\chi^2_{\delta} > \chi^2_{\delta^0}$, тобто χ^2_{δ} попадає в критичну ділянку. Це означає, що розходження між емпіричними і теоретичними частотами суттєве і його не можна пояснити випадковими коливаннями вибірових даних. В такому випадку нульова гіпотеза відхиляється;

2. $\chi^2_{\delta} \leq \chi^2_{\delta^0}$, тобто вирахований критерій не перевищує максимально можливу величину розходжень емпіричних і теоретичних частот, яка може виникнути в силу випадкових коливань вибірових даних.

Табличне значення критерію К. Пірсона визначається при фіксованому рівні значущості і відповідному числі ступенів вільності. Число ступенів вільності дорівнює $k = m - r$, де m – число груп; r – число обернених зв'язків.

Критерій згоди А.Н. Колмогорова (λ) оцінює близькість фактичного розподілу до теоретичного шляхом знаходження величини (D), тобто максимальної різниці нагромаджених (кумулятивних) часток або частот фактичного і теоретичного розподілів.

Критерій А.Н. Колмогорова визначають за формулою

$$\lambda = D * \sqrt{n}, \quad (7.9)$$

де D – абсолютна максимальна різниця кумулятивних часток;

$D = \max |S_d - S_{d^*}|$ або частот $D = \max |S_f - S_{f^*}|$ емпіричного і теоретичного розподілів;

n – число спостережень (чисельність одиниць сукупності).

Якщо розподіл задано в частотах, формула матиме вигляд

$$\lambda = \frac{D}{\sqrt{n}}, \quad (7.10)$$

Методику розрахунку цього показника розглянемо на прикладі даних табл. 7.3:

$$\lambda = \frac{D}{\sqrt{100}} = \frac{5}{10} = 0,5.$$

У спеціальній таблиці імовірностей для критерію згоди « λ » знаходимо, що значенню $\lambda = 0,5$ відповідає імовірність 0,9639. Це означає, що з імовірністю 0,9639 можна стверджувати про нормальний розподіл господарств за врожайністю озимої пшениці.

Таблиця 7.3 – Умовні дані для розрахунку критерію згоди

Номер групи	Нагромаджені частоти		Відхилення $ S_f - S_{f'} $
	Емпіричні S_f	Теоретичні $S_{f'}$	
1	4	2	2
2	11	13	2
3	39	38	1
4	74	69	5
5	90	90	0
6	96	98	2
7	100	100	0

Критерій згоди В.І. Романовського (R) також використовується для оцінки наближення фактичного розподілу до теоретичного. Він визначається за формулою:

$$R = \frac{\chi^2 - k}{\sqrt{2k}}, \quad (7.11)$$

Скориставшись розрахунками прикладу за даними табл. 7.3, визначаємо критерій згоди В.І. Романовського:

$$R = \frac{\chi^2 - k}{\sqrt{2k}} = \frac{8,02 - 4}{\sqrt{2 * 4}} = \frac{4,02}{2,8284} = 1,42 < 3.$$

Якщо при дослідженні наближення фактичного розподілу до теоретичного величина цього виразу менша трьох ($1,42 < 3$), це дає підставу для

ствердження про можливість розподілу за законом даного розподілу. Тобто розподіл господарств за врожайністю озимої пшениці є нормальним.

Критерій згоди П.С. Ястремського (L) використовується для прямої відповіді на питання про міру розбіжності між фактичним і теоретичним розподілом. Для визначення критерію П.С. Ястремського використовується критерій К. Пірсона. В загальному вигляді критерій П.С. Ястремського має вигляд:

$$L = \frac{\chi^2 - n}{\sqrt{2n + 4Q}}, \quad (7.12)$$

де $Q = \frac{(f_i - f'_i)^2}{f_i(1 - P_i)}$ – при кількості груп менша 20 ($n < 20$) та дорівнює 0,6.

За даними попередніх прикладів покажемо розрахунок цього показника:

$$L = \frac{\chi^2 - n}{\sqrt{2n + 4Q}} = \frac{8,02 - 7}{\sqrt{2 * 7 + 4 * 0,6}} = \frac{1,02}{4,02} = 0,25.$$

Так як величина $L < 3$ ($0,25 < 3$), то з ймовірністю 0,997 можна стверджувати, що розподіл господарств за врожайністю озимої пшениці є нормальним.

Всі розглянуті нами критерії згоди дають загальну оцінку ступеня відхилення емпіричного розподілу від нормального, але не визначають його характеру, а тому при суттєвих їх відхиленнях аналіз розподілу доцільно доповнювати характеристиками асиметрії і ексцесу.

Таким чином, для перевірки висунутої гіпотези про відповідність чи невідповідність теоретичного закону розподілу емпіричному можна використати будь-який з наведених критеріїв, які забезпечують дослідження законів розподілу з різною точністю, надійністю і трудоемкістю.

4. Перевірка гіпотези про величину центру розподілу

Припустимо, що ми маємо дані про витрачання повного виду сировини на одиницю продукції до і після модернізації устаткування. До модернізації устаткування таке витрачання становило 30 кг, після модернізації – 28 кг. Середня помилка вибірки 1 кг. Потрібно встановити, чи привела модернізація устаткування до зниження витрат сировини на одиницю продукції.

Нульова гіпотеза заключається в тому, що виробництво продукції до після модернізації устаткування з точки зору впливу на матеріалоємкість суттєво не змінилось, тобто що між генеральними середніми до і після модернізації устаткування немає суттєвої різниці. Тобто нульова гіпотеза означає, що $\bar{z}_0 = \bar{z}_1$, де \bar{z}_0 і \bar{z}_1 – середнє витрачання сировини на одиницю продукції відповідно до і після модернізації устаткування.

Альтернативна гіпотеза в даному випадку може бути сформульована двояко.

1. Модернізація устаткування веде до зміни витрачання сировини на одиницю продукції, тобто нульова гіпотеза заключається в тому, що $\bar{z}_1 \neq \bar{z}_0$. Прийmemo рішення значущості $\alpha = 0,5$, тоді $P(|\bar{z}_1 - \bar{z}_0| \geq t_{\mu\bar{z}}) \geq 0,05$ і критична ділянка задається нерівністю $|\bar{z}_1 - \bar{z}_0| > t_{\mu\bar{z}}$. За таблицями інтегральної функції Лапласа використовуємо коефіцієнт довір'я $t_{\mu\bar{z}} = 1,96$. Таким чином, розмір граничного розходження двох середніх з ймовірністю 0,95 не повинен перевищувати $t_{\mu\bar{z}} = 1,96(1.96 * 100)$. Отже, з ймовірністю 0,95 можна стверджувати, що довірчі межі для генеральної середньої дорівнюватимуть: $28,04 \leq \bar{z}_1 \leq 31,96$. Середнє витрачання матеріалів після модернізації устаткування складає 28 кг і попадає в критичну ділянку. Дані спостереження не співпадають з висунутою нульовою гіпотезою про те,

що між виробництвом продукції до і після модернізації устаткування відсутні суттєві розбіжності з точки зору їх впливу на матеріалоємкість.

2. Модернізація устаткування веде до зниження витрат сировини на виробництво одиниці продукції, тобто нульова гіпотеза заключається в тому, що $\bar{z}_1 < \bar{z}_0$. В цьому випадку розглядається ділянка великих від'ємних відхилень, тобто при $\alpha = 0,05$ ($\bar{z}_1 < \bar{z}_0 - t_{\mu\bar{z}} = 0,05$). В цьому варіанті критична ділянка визначається нерівністю $\bar{z}_1 < \bar{z}_0 - t_{\mu\bar{z}}$. Нульова гіпотеза не буде відхилятися, якщо середні витрати матеріалу на одиницю продукції будуть більші за 28,36 кг ($30 - 1 \cdot 1,64$). Після модернізації устаткування витрачання сировини на одиницю продукції склало 28 кг, тобто з ймовірністю 0,995 можна стверджувати, що нульова гіпотеза повинна бути відхилена і що модернізація устаткування привела до зниження витрачання сировини на виготовлення продукції.

Якщо гіпотеза про величину центру розподілу перевіряється за результатами малої вибірки, то потрібно врахувати, що відношення різниці середніх до стандартної помилки вибірки $t = \frac{\bar{x}_1 - \bar{x}_2}{\mu_{\bar{x}}}$ має розподіл Стюдента з $n - 1$ ступенями волі. Припустимо, що дані про витрачання сировини на одиницю продукції були отримані за результатами перевірки 17 виробів. За результатами спостережень значення емпіричного середнього кореляційного відношення склало 5,0. Обчислимо величину t-критерію за формулою

$$t_{\delta} = \frac{\bar{z}_1 - \bar{z}_2}{\frac{S}{\sqrt{n-1}}} = \frac{|28 - 30| \cdot 4}{5} = 1,6.$$

За таблицею розподілу Стюдента значення t -критерію для числа ступенів свободи $(n - 1) = 16$ і рівня значущості 0,05 дорівнює $t_T = 2,12$. Так як фактичне значення $[t_{\delta} < t_T (1,6 < 2,12)]$ не перевищує табличне, то нульова гіпотеза не відхиляється, тобто у нас відсутні достатні підстави

вважати, що модернізація устаткування веде до зниження матеріалоємкості.



Питання для самоконтролю по темі 7

1. *Поняття про статистичну гіпотезу.*
2. *Можливі помилки при перевірці гіпотез.*
3. *Завдання для перевірки певних статистичних гіпотез.*
4. *Визначення понять „рівень значущості” і „критична ділянка”.*
5. *Визначення понять „ділянки відхилень”.*
6. *Причини виникнення спостережень, які відрізняються від загального рівня.*
7. *Критерій Ф. Груббса.*
8. *Критерій К.Пірсона.*
9. *Критерій згоди А.Н. Колмогорова.*
10. *Критерій згоди В.І. Романовського.*
11. *Критерій згоди П.С. Ястремського.*
12. *Гіпотеза про величину центру розподілу.*

ТЕМА 8. СТАТИСТИЧНІ МЕТОДИ АНАЛІЗУ КОРЕЛЯЦІЙНИХ ЗВ'ЯЗКІВ

Питання для теоретичної підготовки

1. Види взаємозв'язку показників соціально-економічних явищ.
2. Загальні методи вивчення зв'язків.
3. Кореляційний і регресійний методи аналізу зв'язків. Парна кореляційна залежність.
4. Визначення тісноти залежності в кореляційно-регресійному аналізі.
5. Багатофакторний кореляційно-регресійний аналіз. Множинна регресія.

1. Види взаємозв'язку показників соціально-економічних явищ

Хоч зв'язки між різними явищами в реальному світі складні й різноманітні, однак їх можна певним чином класифікувати. Розкриваючи взаємозв'язки і взаємозалежності між явищами можна пізнати їх суть і закони розвитку. Тому вивчення взаємозв'язків є основним завданням всякого статистичного аналізу.

Суспільні явища або окремі їх ознаки, які впливають на інші й обумовлюють їх зміну, називаються **факторними**, а суспільні явища або окремі їх ознаки, які змінюються під впливом факторних, називаються **результативними**.

У техніці й природознавстві часто мова йде про **функціональну залежність** між змінними X і Y , коли кожному можливому значенню X поставлене в однозначну відповідність певне значення Y . Такі зв'язки пізнаються за допомогою експериментів, виражаються у формі законів.

Соціально-економічні явища відбуваються в обстановці дії численних факторів, вплив кожного з яких може бути яким завгодно малим, а їх кількість – якою завгодно великою. У цих випадках зв'язок втрачає свою строгу функціональність і досліджуване явище переходить не у певний стан, а у

один з можливих станів. У цьому випадку для визначення зв'язків замість експерименту використовують певні статистичні методи.

Така залежність називається **стохастичною** і полягає вона у тому, що змінна Y перебуває в стохастичній залежності від X , якщо кожному значенню X відповідає ряд розподілу Y і зі зміною X ці ряди закономірно змінюються. Якщо ж ці ряди не змінюються або змінюються випадково, то такої залежності не існує.

Частковим випадком стохастичної залежності у практиці досліджень є **статистична залежність**, коли умовне математичне очікування однієї випадкової змінної є функцією значення, прийнятого іншою випадковою змінною, тобто

$$M_0 (Y / x) = f (x), \quad (8.1)$$

де M_0 – показник середнього значення, прийнятого випадковою змінною, який дорівнює середньозваженій всіх можливих значень змінної, у якій вагами є ймовірності відповідних подій.

Знання статистичної залежності між випадковими змінними дозволяє прогнозувати значення залежної змінної у припущенні, що незалежна змінна прийме певне значення.

Спрощуючи судження й уникаючи помилок прогнозу на практиці при вивченні статистичної залежності, прийнято переходити від умовного математичного очікування випадкової змінної до умовного середнього значення цієї змінної.

В цьому випадку залежність між однією випадковою змінною й умовним середнім значенням іншої випадкової змінної називають **кореляційною залежністю**.

$$M_0 (Y / X = x) = y (x), \quad (8.2)$$

де $M_0 (Y / X = x)$ – математичне очікування випадкової змінної Y за умови, що випадкова величина X прийняла значення x .

Основоположниками теорії кореляції слід вважати англійських біометриків Френсіса Гальтона (1822 - 1911) і Карла Пірсона (1857 - 1936).



Кореляція (від лат. *correlatio* – взаємозв'язок) – в статистиці це поняття, яке відображає наявність зв'язку між явищами, процесами і величинами, що їх характеризують. **Кореляція** – це систематичний і обумовлений зв'язок між двома рядами даних. **Кореляція** – зв'язок змінних, при якому одному значенню однієї ознаки відповідає кілька значень іншої ознаки, що відхиляється у той чи інший бік від свого середнього значення.

Основним завданням вивчення кореляційних залежностей слід визнати виявлення причин зміни досліджуваного явища, події або факту. Причому розрізняють **факторну ознаку** як *ознаку-причину* й **результативну ознаку** як *ознаку-слідство*.

Одній й тій же величині факторної ознаки відповідають різні значення результативної ознаки, що утворюють ряд розподілу. Кореляційна залежність проявляється лише в середньому, у масі випадків.

За напрямом розрізняють залежності прямі й зворотні, **за аналітичним вираженням** – лінійні й нелінійні. При прямому взаємозв'язку зі збільшенням значення факторної ознаки спостерігається тенденція до збільшення індивідуальних і середніх значень результативної ознаки. При зворотній залежності із зростанням факторної ознаки значення результативної ознаки зменшуються. Існують також безпосередні й непрямі залежності. Фактор X може впливати на Y безпосередньо або побічно, через інший фактор W .

Залежність двох ознак називається **парною кореляцією**, вплив декількох факторів на результативну ознаку – **багатофакторним** (множинним).

При вивченні кореляційних залежностей між показниками-ознаками вирішуються наступні завдання:

- попередній аналіз властивостей сукупності;

- встановлення факту наявності залежності, визначення її напрямку й форми залежності;
- вимір ступеня тісноти залежності між ознаками;
- побудова регресійної моделі, тобто знаходження аналітичного вираження форми зв'язку;
- оцінка адекватності моделі, її логічна й економічна інтерпретація.

2. Загальні методи вивчення зв'язків

Зв'язки і залежності суспільних явищ вивчаються різними методами, які дають уявлення про їх наявність і характер. До цих методів відносять:

- балансовий метод;
- метод порівняння паралельних рядів;
- графічний метод;
- метод аналітичних групувань;
- індексний метод;
- кореляційно-регресійний аналіз тощо.

Одним з найпоширеніших методів статистичного вивчення зв'язків суспільних явищ є **балансовий метод** як прийом аналізу зв'язків і пропорцій в економіці.

Статистичний баланс являє собою систему показників, яка складається із двох сум абсолютних величин, пов'язаних між собою знаком рівності.

$$A + B = C + D$$

Цю балансову ув'язку можна зобразити через **балансове рівняння**:

$$\sum_{i=1}^n x_i + \sum_{j=1}^m y_j - \sum_{k=1}^p z_k = \sum_{l=1}^q w_l$$

Наведена балансова рівність характеризує єдиний процес руху матеріальних ресурсів і показує взаємозв'язок і пропорції окремих елементів цього процесу.

Метод порівняння паралельних рядів полягає в тому, що отримані у результаті групування і лічильної обробки матеріали статистичного спостереження ранжуються паралельними рядами за факторною ознакою.

Паралельно записуються значення результативної ознаки. Це дає можливість, порівнюючи їх, простежити співвідношення, виявити існування зв'язку і його напрямку.

Коефіцієнт Фехнера оцінює силу зв'язку на основі порівняння знаків відхилень значень варіантів від їх середньої по кожній ознаці.

Знак мінус означає, що значення ознаки менше середньої, а знак плюс – більше середньої. Співпадіння знаків за обома ознаками означає узгоджену варіацію, неспівпадіння – порушення такої узгодженості. За цим принципом побудований коефіцієнт Фехнера:

$$K_{\phi} = \frac{\sum C - \sum H}{\sum C + \sum H}, \quad (8.3)$$

де $\sum C$ – сума знаків, які співпали по обох рядах;

$\sum H$ – сума знаків, які не співпали.

Коефіцієнт Фехнера коливається в межах від « + 1 » до « - 1 ». При наближенні цього коефіцієнта до « + 1 » спостерігається пряма і сильна узгодженість, при « - 1 » маємо сильну але обернену узгодженість. При нулю узгодженість між досліджуваними ознаками відсутня.

Графічний метод виявлення кореляційної залежності полягає у зображенні на графіку статистичних характеристик, отриманих в результаті зведення і обробки вихідної інформації, яке наочно показує форму зв'язку між досліджуваними ознаками та його напрямком.

Приклад. Зв'язок між вартістю основних виробничих фондів і випуском продукції можна наочно уявити, якщо побудувати графік. Позначивши на графіку точки, які відповідають значенням « X » і « Y », отримаємо кореляційне поле, де за характером розміщення точок можна судити про напрямку і силу зв'язку.

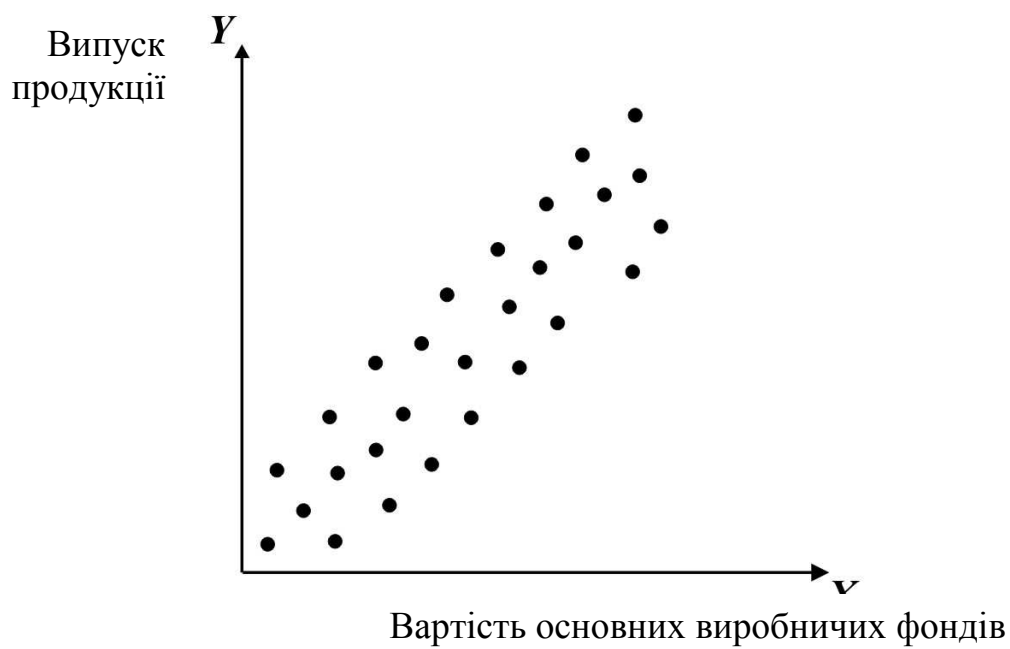


Рис. 8.1 – Кореляційне поле зв'язку між вартістю основних виробничих фондів і випуском продукції

Якщо точки розташовані хаотично по усьому полю, це говорить про відсутність залежності між двома ознаками; якщо вони сконцентровані навколо осі, яка йде від нижнього лівого кута до верхнього правого – це пряма залежність між досліджуваними ознаками; якщо точки будуть сконцентровані навколо осі, яка проляже від верхнього лівого кута до нижнього правого – маємо обернену залежність.

Метод статистичних групувань як прийом виявлення кореляційної залежності відноситься до числа найважливіших прийомів дослідження взаємозв'язків. Для виявлення залежності між ознаками за допомогою цього методу матеріал статистичного спостереження групується за факторною ознакою; для кожної групи вираховуються середні значення як факторної так і результативної ознаки. Порівнюючи зміни середніх значень результативної ознаки в міру зміни середніх значень факторної ознаки, виявляють характер зв'язку між ними.

Статистичні групування, проведені з метою виявлення і аналізу взаємозв'язків між ознаками, називаються *аналітичними*.

Групування дозволяє також виявити одночасний вплив декількох чинників на результативну ознаку. Для цього проводять *комбіновані* групування, дані яких викладають в комбінованих таблицях.

Аналітичні групування характеризують лише загальні риси зв'язку, його тенденцію, але не дають кількісної оцінки його сили. На основі аналітичних групувань це завдання розв'язується за допомогою розрахунку емпіричного кореляційного відношення.

Для кількісної оцінки зв'язку між явищами на базі матеріалів аналітичного групування вираховують коефіцієнт детермінації і емпіричне кореляційне відношення.

Коефіцієнт детермінації показує ступінь варіації ознаки під впливом чинника покладеного в основу групування.

Він визначається як відношення міжгрупової дисперсії до загальної.

3. Кореляційний і регресійний методи аналізу зв'язків

Кореляційний і регресійний методи аналізу зв'язків покликані вирішити **два основні завдання**:

1. визначення за допомогою рівнянь регресії аналітичної форми залежності між варіацією ознак X і Y ;
2. встановлення ступеню тісноти залежності між ознаками (якою мірою варіація X спричиняється варіацію Y).

В основу виявлення і встановлення аналітичної форми залежності покладене застосування в аналізі вихідної інформації математичних функцій. При вивченні зв'язку показників застосовують різного виду рівняння прямолінійної й криволінійної залежностей.

У випадку лінійної форми зв'язку результативна ознака змінюється під впливом факторної ознаки рівномірно. При аналізі прямолінійної залежності застосовується рівняння прямої:

$$Y_x = a_0 + a_1 x, \quad (8.4)$$

де Y_x – вирівняне середнє значення результативної ознаки;

a_0 – значення Y при $x = 0$;

a_1 – коефіцієнт регресії;

Коефіцієнт регресії a_1 показує, на скільки зміниться результативна ознака Y при зміні факторної ознаки x на одиницю. Якщо a_1 має позитивний знак, то зв'язок прямий, якщо від'ємний – то зв'язок обернений.

Дослідження форми зв'язку іноді приводить до необхідності використання нелінійних рівнянь регресії. При аналізі криволінійної залежності застосовується формули ряду математичних функцій:

■ напівлогарифмічна: $y_x = a_0 + a_1 \lg x, \quad (8.5)$

■ показова: $y_x = a_0 + a_1^x, \quad (8.6)$

■ степенева: $y_x = a_0 x^{a_1}, \quad (8.7)$

■ параболічна: $y_x = a_0 + a_1 x + a_2 x^2, \quad (8.8)$

■ гіперболічна: $y_x = a_0 + a_1 * 1/x, \quad (8.9)$

та інші.

Після вибору форми залежності, використовуючи вихідні дані, необхідно визначити параметри кореляційного рівняння, тобто a_0 і a_1, a_2, \dots . Для їх знаходження існують різні методи:

- метод середніх значень;
- метод кінцевих різниць;
- метод найменших квадратів.

Після знаходження параметрів рівнянь будують синтезовану *трендову* модель (фактичну).

Найчастіше це здійснюється способом вирівнювання емпіричних (факторних) даних **методом найменших квадратів**. Якщо спостереження нанести на графік, що характеризує їх розсіювання, то лінія, що представляє це рівняння, буде розташована так, що сума квадратів відстаней по вертикалі між крапками й цією лінією буде мінімальною.

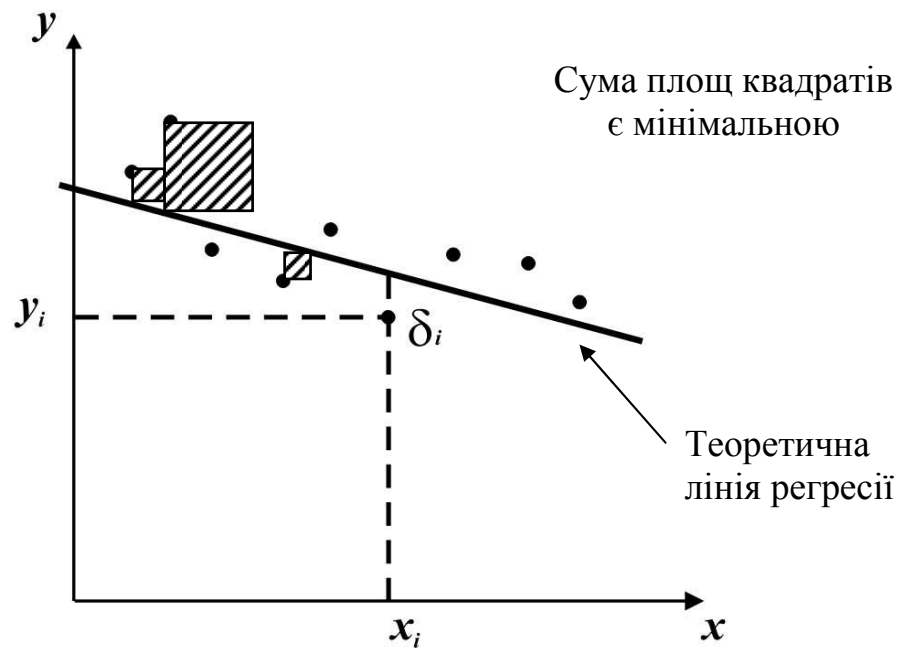


Рис. 8.2 – Метод найменших квадратів

В основу цього методу покладена вимога мінімальності сум квадратів відхилень емпіричних даних y_i від теоретичних (вирівняних) y_{xi} :

$$\sum (y_i - \bar{y}_{xi})^2 = \min, \quad (8.10)$$

Це дає можливість отримати найкращі оцінки параметрів a_0 і a_1 . Для їх обчислення треба скласти і розв'язати систему нормальних рівнянь. Лінійній моделі відповідає система двох нормальних рівнянь з двома невідомими:

$$\begin{cases} \sum y_i = na_0 + a_1 \sum x_i \\ \sum x_i y_i = a_0 \sum x_i + a_1 \sum x_i^2, \end{cases} \quad (8.11)$$

де n – число членів в кожному з двох порівнювальних рядів;

$\sum x_i$ – сума значень факторної ознаки;

$\sum x_i^2$ – сума квадратів значень факторної ознаки;

$\sum y_i$ – сума значень результативної ознаки;

$\sum x_i y_i$ – сума добутків значень факторної ознаки на значення результативної ознаки.

Для вирішення системи застосовують спосіб визначників, який дозволяє зводити до мінімуму неточності округлень у розрахунках параметрів рівнянь регресії:

$$a_0 = \frac{\sum y_i \sum x_i^2 - \sum x_i y_i \sum x_i}{n \sum x_i^2 - \sum x_i \sum x_i}, \quad (8.12)$$

$$a_1 = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{n \sum x_i^2 - \sum x_i \sum x_i}, \quad (8.13)$$

Визначивши параметри і підставляючи в рівняння регресії (8.3) послідовно значення факторної ознаки (x_i), ми отримаємо вирівняні значення результативної ознаки (y_{xi}), які покажуть, яким чином вона теоретично розподіляється.

При статистичному аналізі криволінійного зв'язку часто застосовують напівлогарифмічну функцію:

$$y_x = a_0 + a_1 \lg x, \quad (8.14)$$

Параметри рівняння визначаються із системи нормальних рівнянь, що відповідають вимогам методу найменших квадратів:

$$\begin{cases} na_0 + a_1 \sum \lg x_i = \sum y_i \\ a_0 \sum \lg x_i + a_1 \sum (\lg x_i)^2 = \sum y_i * \lg x_i \end{cases}, \quad (8.15)$$

З використанням методу визначників складаються алгоритми розрахунку параметрів рівняння:

$$a_0 = \frac{\sum y_i \sum (\lg x_i)^2 - \sum y_i \lg x_i \sum \lg x_i}{n \sum (\lg x_i)^2 - \sum \lg x_i \sum \lg x_i}, \quad (8.16)$$

$$a_1 = \frac{n \sum y_i \lg x_i - \sum y_i \sum \lg x_i}{n \sum (\lg x_i)^2 - \sum \lg x_i \sum \lg x_i}, \quad (8.17)$$

При використанні нелінійних рівнянь регресії визначення їх параметрів методом найменших квадратів здійснюють приведенням їх до лінійного виду певними перетвореннями:

- 1) степенева $y_x = a_0 x^{a_1}$ приводиться до лінійного вигляду логарифмуванням $\lg y = \lg a_0 + a_1 \lg x$; Подальші розрахунки аналогічні лінійній моделі;
- 2) гіперболу $y_x = a_0 + a_1 * 1/x$ приводять до лінійного вигляду, змінивши x новою змінною (її зворотнім значенням $z = 1/x$), а далі вирішуючи рівняння: $y_x = a_0 + a_1 z$;
- 3) парабола другого порядку $y_x = a_0 + a_1 x + a_2 x^2$. Якщо замінити квадрат значень факторної ознаки ($z = x^2$), то ми дістанемо лінійну функцію від двох змінних: $y_x = a_0 + a_1 x + a_2 z$.

Для економічної інтерпретації лінійних і нелінійних зв'язків між двома досліджуваними явищами використовують розраховані на підставі рівнянь регресії коефіцієнти еластичності.

Коефіцієнт еластичності показує, на скільки відсотків зміниться в середньому результативна ознака \bar{y} при зміні факторної ознаки \bar{x} на 1%.

Для лінійної залежності коефіцієнт еластичності визначається за формулою:

$$\varepsilon = a_1 * \frac{\bar{x}}{\bar{y}}, \quad (8.18)$$

де ε – коефіцієнт еластичності.

Якщо залежність між ознаками представлена за даними, вирівняними за параболою другого порядку, то коефіцієнт еластичності визначається за формулою:

$$\varepsilon = (a_1 + a_2 x) * \frac{\bar{x}}{\bar{y}}, \quad (8.19)$$

4. Визначення тісноти залежності в кореляційно-регресійному аналізі

Визначення тісноти зв'язку (залежності) в кореляційно-регресійному аналізі, як і в методі аналітичного групування, ґрунтується на правилі складання дисперсій. На відміну від аналітичного групування, де оцінками лінії регресії є групові середні результативної ознаки, в кореляційно-регресійному аналізі – теоретичні значення результативної ознаки.

Для статистичної оцінки тісноти зв'язку застосовуються наступні показники:

- **Загальна дисперсія** результативної ознаки σ_y^2 , що відображає сукупний вплив всіх факторів:

$$\sigma_y^2 = \frac{\sum (y_i - \bar{y})^2}{n}, \quad (8.20)$$

де відхилення $y_i - \bar{y}$ обумовлені тим, що сполучення впливу факторів на варіацію ознаки y , для кожної одиниці аналізованої сукупності різний.

- **Факторна дисперсія** результативної ознаки $\sigma_{y_x}^2$, що відображає варіацію результативної ознаки y тільки від впливу досліджуваного фактору x :

$$\sigma_{y_x}^2 = \frac{\sum (y_{x_i} - \bar{y})^2}{n}, \quad (8.21)$$

де відхилення $y_{x_i} - \bar{y}$ характеризують коливання вирівняних значень y_{x_i} від їх загальної середньої величини \bar{y} .

■ **Залишкова дисперсія** σ_ε^2 , що відображає варіацію результативної ознаки y від всіх інших факторів, крім x :

$$\sigma_\varepsilon^2 = \frac{\sum (y_i - y_{x_i})^2}{n}, \quad (8.22)$$

де відхилення $y_i - y_{x_i}$ характеризують коливання емпіричних (фактичних) значень результативної ознаки y від їх вирівняних значень y_{x_i} .

Для пояснення останнього твердження відзначимо, що відхилення y_i від \bar{y} (та відповідних дисперсій) можна подати у вигляді:

$$y_i - \bar{y} = (y_i - y_{x_i}) + (y_{x_i} - \bar{y}), \quad (8.23)$$

Графічну ілюстрацію цього рівняння можна навести в наступному вигляді:

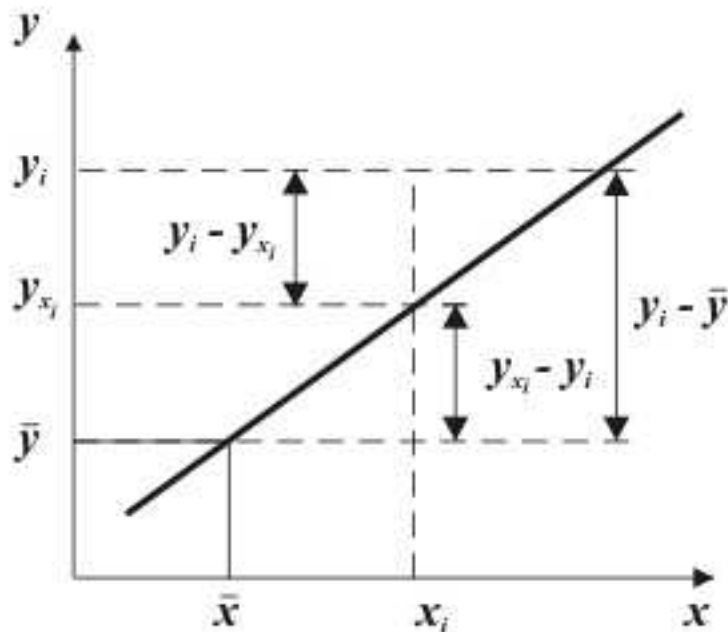


Рис. 8.3 – Варіація результативної ознаки

Різниця $(y_{x_i} - \bar{y})$ утворюється залежно від різниці $(x_i - \bar{x})$, тобто варіація результативної ознаки обумовлюється факторною ознакою.

Друга частина різниці $y_i - \bar{y}$ є різницею між y_i та її оцінкою y_{x_i} . Ця різниця є помилкою моделі, до неї входить вплив неврахованих факторів.

Співвідношення між факторною й загальною дисперсіями характеризує **міру тісноти зв'язку** між ознаками x и y :

$$\frac{\sigma_{y_x}^2}{\sigma_y^2} = R^2, \quad (8.24)$$

Показник R^2 називається **індексом детермінації** (причинності). Він виражає частку факторної дисперсії в загальній дисперсії, тобто характеризує, яка частина загальної варіації результативної ознаки y пояснюється досліджуваним фактором x , тобто відповідає лінійному рівнянню регресії.

На підставі наведеної формули визначається **індекс кореляції R** :

$$R = \sqrt{\frac{\sigma_y^2 - \sigma_\varepsilon^2}{\sigma_y^2}} \quad \text{або} \quad R = \sqrt{\frac{\sigma_{y_x}^2}{\sigma_y^2}}, \quad (8.25)$$

Індекс кореляції приймає значення від « 0 » до « 1 ». Коли $R = 0$, то зв'язку між варіацією ознак y та x немає. Залишкова дисперсія дорівнює загальній ($\sigma_\varepsilon^2 = \sigma_y^2$), а теоретична дорівнює нулю ($\sigma_{y_x}^2 = 0$). Всі теоретичні значення y_x збігаються з середнім значенням \bar{y} .

При $R = 1$ теоретична дисперсія дорівнює загальній ($\sigma_{y_x}^2 = \sigma_y^2$), а залишкова – ($\sigma_\varepsilon^2 = 0$). Фактичні значення збігаються з теоретичними, тобто зв'язок між ознаками є функціональним.

Індекс кореляції придатний для вимірювання тісноти зв'язку при будь-якій її формі. Напрямок зв'язку він не показує.

Для вимірювання тісноти зв'язку і визначення його напрямку при лінійній залежності визначається **лінійний коефіцієнт кореляції r** за формулою:

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{n\sigma_x \sigma_y}, \quad (8.26)$$

На практиці лінійний коефіцієнт кореляції визначається за формулою:

$$r = \frac{\overline{xy} - \bar{x} * \bar{y}}{\sigma_x \sigma_y}, \quad (8.27)$$

Значення r коливається в межах від « -1 » до « $+1$ ». Додатне значення r вказує на прямий зв'язок між ознаками, а від'ємне – на зворотній.

Для одержання висновків про практичну значимість синтезованих в аналізі моделей показанням тісноти зв'язку дається якісна оцінка. Це здійснюється на підставі наступної приблизної схеми (табл. 8.1):

Таблиця 8.1 – Шкала ступеня зв'язку

Показник тісноти зв'язку r	0,1 – 0,3	0,3 – 0,5	0,5 – 0,7	0,7 – 0,9	0,9 – 0,99
Характеристика ступеню зв'язку	слабка	помірна	помітна	висока	досить висока

5. Багатофакторний кореляційно-регресійний аналіз.

Множинна регресія

Наведений вище аналіз статистичних сукупностей дозволяє вивчати взаємозв'язок тільки двох змінних.

На практиці в багатьох випадках на результативну ознаку впливає не один, а декілька чинників (факторів). Між ними існують складні взаємозв'язки, тому їх вплив на результативну ознаку комплексний і його не можна розглядати як просту суму ізольованих впливів.

Багатофакторний кореляційно-регресійний аналіз дозволяє оцінити міру впливу на досліджуваний результативний показник кожного із введених в модель факторів при зафіксованому на середньому рівні інших факторів. При цьому важливою умовою тут є відсутність функціонального зв'язку між факторами.

Математично завдання зводиться до знаходження аналітичного виразу, який найкращим чином відображав зв'язок результативної ознаки з факторними, тобто описував функцію:

$$\hat{y} = f(x_1, x_2, \dots, x_m), \quad (8.28)$$

Форму зв'язку можна визначити шляхом перебору функцій різних типів, але це пов'язане з великою кількістю зайвих розрахунків. Раніше ми довели, що будь-яку функцію багатьох змінних шляхом логарифмування або заміни змінних можна звести до лінійного виду.

У цьому випадку статистична модель може бути представлена рівнянням регресії з декількома змінними величинами. Така регресія називається **множинною**.

Наприклад, лінійна регресія з m незалежними змінними матиме вигляд:

$$\hat{y}_x = a_0 x_0 + a_1 x_1 + a_2 x_2 + \dots + a_m x_m, \quad (8.29)$$

де \hat{y}_x – розраховані значення результативної ознаки-функції;

x_1, x_2, \dots, x_m – значення факторних ознак;

$a_0, a_1, a_2, \dots, a_m$ – параметри рівняння.

При оцінці параметрів цього рівняння в кожному i -му спостереженні фіксують значення результативної ознаки в i -й факторній ознаці $x_{i_0} \dots x_{i_m}$.

Для обчислення параметрів складається і розв'язується система нормальних рівнянь. Наприклад, для лінійної двохфакторної моделі це буде система трьох рівнянь:

$$\begin{cases} na_0 + a_1 \sum x_1 + a_2 \sum x_2 = \sum y \\ a_0 \sum x_1 + a_1 \sum x_1^2 + a_2 \sum x_1 x_2 = \sum x_1 y \\ a_0 \sum x_2 + a_1 \sum x_1 x_2 + a_2 \sum x_2^2 = \sum x_2 y \end{cases}, \quad (8.30)$$

Оцінки параметрів рівняння регресії за допомогою методу найменших квадратів у випадку множинної регресії зручніше обчислювати з використанням стандартних програм і ПК.

Сукупний коефіцієнт множинної кореляції r_y характеризує тісноту зв'язку результативної y в i факторних (x_1, x_2, \dots, x_m) ознак і в загальному випадку визначають за формулою:

$$r_y = \sqrt{\frac{\sigma_{y12\dots m}^2}{\sigma_{y^2}}} = \sqrt{1 - \frac{\sigma_{y(12\dots m)}^2}{\sigma_{y^2}}}, \quad (8.31)$$

значення коефіцієнта перебуває в межах $0 \leq r_y \leq 1$

де $\sigma_{y12\dots m}^2$ – факторна дисперсія;
 $\sigma_{y(12\dots m)}^2$ – залишкова дисперсія;
 σ_y^2 – дисперсія результативної ознаки.

$$\sigma_{y12\dots m}^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{n-1}, \quad (8.32)$$

$$\sigma_{y(12\dots m)}^2 = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n-1}, \quad (8.33)$$

$$\sigma_y^2 = \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n-1}, \quad (8.34)$$

де \hat{y}_i – розрахункове значення результативної ознаки;
 \bar{y} – середнє значення результативної ознаки.

$$B_y = r_y^2 = \frac{\sigma_{y12\dots m}^2}{\sigma_{y^2}}, \quad (8.35)$$

Квадрат величини r_y є *коефіцієнтом множинної детермінації* B_y й характеризує частку впливу обраних ознак на результативний фактор.



Питання для самоконтролю по темі 8

1. *Поняття про взаємозв'язки суспільно-економічних явищ.*
2. *Поняття про факторні і результативні ознаки.*
3. *Поняття про стохастичний зв'язок.*
4. *Відмінність кореляційної та функціональної залежності.*
5. *Поняття про форми та напрямки залежності.*
6. *Поняття про статистичний баланс.*
7. *Сутність методу паралельних рядів.*
8. *Оцінка зв'язку за коефіцієнтом Фехнера.*
9. *Графічний метод виявлення кореляційних зв'язків.*
10. *Поняття про метод статистичних групувань.*
11. *Поняття про кореляційно-регресійний аналіз.*
12. *Основні завдання кореляційно-регресійний аналізу.*
13. *Основні етапи проведення кореляційно-регресійний аналізу.*
14. *Сутність коефіцієнтів регресії.*
15. *Для чого використовують коефіцієнт регресії?*
16. *Поняття про нелінійні залежності.*
17. *Сутність методу найменших квадратів.*
18. *Система нормальних рівнянь для знаходження коефіцієнтів регресії.*
19. *Багатофакторний кореляційно-регресійний аналіз.*
20. *Множинна регресія.*
21. *Сукупний коефіцієнт множинної кореляції.*
22. *Методи приведення нелінійних залежностей до лінійних.*

ТЕМА 9. АНАЛІЗ ІНТЕНСИВНОСТІ ДИНАМІКИ

Питання для теоретичної підготовки

1. Визначення рядів динаміки. Види рядів динаміки.
2. Правила побудови динамічних рядів.
3. Основні статистичні характеристики рядів динаміки.
4. Вибір бази порівняння і середні темпи рядів динаміки.
5. Середні показники динаміки.

1. Визначення рядів динаміки. Види рядів динаміки

У практиці статистичної роботи доволі часто мають справу з великою кількістю різноманітних показників, які характеризують розвиток вивчаємих явищ у часі. Для усвідомлення і аналізу досліджуваних статистичних даних, їх систематизують, будуючи хронологічні ряди, які називаються **рядами динаміки**.

Ряди динаміки – це послідовність статистичних показників, які характеризують закономірності й особливості зміни масових суспільних явищ і процесів в часі.

Ряди динаміки складаються з рівнів ряду. **Рівні ряду (y)** – це показники динамічного ряду. Рівні рядів динаміки можуть бути виражені абсолютними або відносними величинами.

Будь-який ряд динаміки складається з двох основних елементів:

- 1) *періодів або моментів* часу, до яких відносяться рівні ряду (t);
- 2) *статистичних показників*, які характеризують рівні ряду (y).

Залежно від характеру рівнів ряду розрізняють два види рядів динаміки:

Моментним називається ряд динаміки, величини рівнів якого характеризують стан явищ на певний момент часу. Рівні моментного ряду підсумувати не має змісту.

Інтервальним називається такий ряд динаміки, величини рівнів якого характеризують розміри суспільних явищ за певні періоди часу (день, місяць, квартал тощо). Сума рівнів інтервального ряду динаміки характеризує рівень даних явища за більш тривалий проміжок часу.

Ряди динаміки бувають одномірні і багатомірні.

Одномірні ряди динаміки характеризують зміну одного показника (валовий внутрішній продукт).

Багатомірні ряди динаміки характеризують зміну двох, трьох і більше показників.

У свою чергу багатомірні динамічні ряди поділяються на паралельні ряди і ряди взаємозв'язаних показників.

Паралельні ряди динаміки відображають зміну або одного і того самого показника щодо різних об'єктів (чисельність працюючих по підприємствах), або різних показників щодо одного і того самого об'єкта (середня зарплата, обсяги реалізації продукції, продуктивність праці підприємств міста).

Ряди взаємозв'язаних показників характеризують залежність одного явища від іншого (залежність прибутку підприємства від обсягів реалізації).

За повнотою часу динамічні ряди поділяються на повні і неповні.

В **повних** динамічних рядах дати або періоди йдуть один за одним з рівними інтервалами.

В **неповних** динамічних рядах в послідовності часу спостерігаються нерівні інтервали.

За способом вираження рівнів динамічного ряду вони поділяються на ряди **абсолютних, відносних і середніх** величин.

2. Правила побудови динамічних рядів

При побудові динамічних рядів для дослідження розвитку суспільних явищ в часі дуже важливо дотримуватись правил їх побудови. Важ-

ливим правилом побудови динамічних рядів є вимога порівняльності усіх рівнів ряду між собою. Показники рядів динаміки повинні бути порівняльні за територією, колом охоплюваних об'єктів, способами розрахунків, періодами часу, одиницями виміру.

Важливою вимогою любых динамічних порівнянь є вимога **порівняльності території**, до якої відносяться рівні динамічного ряду. У кожному окремому випадку питання порівняльності розв'язується в залежності від мети дослідження. Для приведення даних динамічного ряду до порівняльного виду проводиться перерахунок попередніх даних з врахуванням нових меж (кордонів).

Статистичні дані, які необхідні для побудови ряду динаміки повинні бути порівняльні за **колом охоплюваних об'єктів**. Непорівняльність може виникнути внаслідок переходу деяких об'єктів із одного підпорядкування у інше. Порівняльність за колом охоплюваних об'єктів забезпечується **зімкненням динамічних рядів** шляхом заміни абсолютних рівнів відносними.

У моментних рядах динаміки може виникати непорівняльність за **критичним моментом реєстрації** рівнів явищ, які піддаються сезонним та добовим коливанням (споживання електроенергії взимку і влітку, вдень і вночі).

Рівні динамічного ряду повинні бути порівняльні за **методикою їх розрахунку**. Наприклад, за попередні роки чисельність робітників заводу була визначена на початок кожного місяця, тобто на певну дату, а в наступні роки – як середньомісячна чисельність.

Статистичні дані динамічного ряду можуть бути непорівняльними за **різними періодами або тривалістю часу**. Наприклад, обсяг виробництва електричної енергії за різні роки потрібно порівнювати тільки таким чином: січень з січнем, грудень з груднем і аналогічно надалі. Інтервали часу, за які наведені дані динамічного ряду, повинні бути рівні (місяць, квартал, півріччя тощо).

Непорівняльність *через різні одиниці виміру* виникає внаслідок того, що ряд явищ обліковується паралельно в двох одиницях виміру. Наприклад, залізобетонні конструкції обліковуються в тоннах і кубічних метрах, електромотори – в штуках і кіловатах потужності. Порівняльність за одиницями виміру вимагає, щоб рівні динамічного ряду завжди були виражені в одних і тих самих одиницях виміру.

Непорівняльність рядів динаміки через одиниці виміру виникає і внаслідок непорівняльності грошової оцінки (змінюються грошові одиниці, вплив інфляційних процесів, змінюється валютний курс тощо). Для приведення до порівняльного виду таких рядів динаміки усі попередні рівні досліджуваних ознак перераховуються за діючою грошовою оцінкою.

Непорівняльність статистичних показників динаміки може бути зумовлена також **різною структурою сукупності** за ряд років. Для приведення даних таких рядів до порівняльного виду використовують так звану стандартизацію структури.

3. Основні статистичні характеристики рядів динаміки

Метою аналізу рядів динаміки є розкриття і характеристика закономірностей, які проявляються на різних етапах розвитку того або іншого соціально-економічного явища, виявлення тенденцій і напрямів його розвитку та їх особливостей.

Зміни показників рядів динаміки можуть бути двох видів: зростання цих показників або їх зменшення. Іноді **зменшення** показників рядів динаміки помилково називають «**темпами зростання**», «**приростами**» тощо. Тому надалі доцільно розрізняти темпи зростання і темпи зменшення, абсолютні прирости і абсолютні зменшення тощо.

Окрім рівнів, середніх величин та показників варіації ряди динаміки характеризують наступні показники:

- 1) абсолютні відхилення (абсолютні прирости та абсолютні зменшення);

- 2) відносні відхилення у відсотках (відносні прирости та відносні зменшення);
- 3) темпи змін (темпи зростання та темпи зменшення);
- 4) абсолютне значення одного відсотка відхилення (приросту або зменшення).

У свою чергу ці показники розподіляють на **базисні** та **ланцюгові**. Базисні показники мають постійну базу порівняння, а ланцюгові – змінну. **База порівняння** – це величина, з якою порівнюють той чи інший показник. Як правило, в базисних показниках базою порівняння є перший рівень ряду динаміки, а в ланцюгових – попередній.

Абсолютне відхилення (Δy_t) – це різниця між двома рівнями ряду динаміки, яка визначається за формулами:

■ для ланцюгових відхилень:

$$\Delta^l y_t = y_t - y_{t-1} , \quad (9.1)$$

де $y_t; y_{t-1}$ - порівнюваний та попередній рівень ряду динаміки;

■ для базисних відхилень:

$$\Delta^b y_t = y_t - y_b , \quad (9.2)$$

де y_b – базисний (перший) рівень .

Абсолютне відхилення показує абсолютну величину зміни рівня ряду динаміки за певний час.

Відносне відхилення ($\Delta_{\%}$) – це відношення абсолютного відхилення до відповідного рівня ряду динаміки у відсотках:

$$\Delta_{\%} = \frac{\Delta_a}{y_i} \cdot 100\% , \quad (9.3)$$

де Δ_a – відповідне абсолютне відхилення (для ланцюгових показників – ланцюгове, для базисних показників – базисне);

y_i – ланцюговий або базисний рівень ряду динаміки.

Темп зміни (T_i) – це відношення двох рівнів ряду динаміки у відсотках:

$$T_i = \frac{y_t}{y_i} \cdot 100\% , \quad (9.4)$$

де y_t - порівнюваний рівень;

y_i - попередній рівень для ланцюгових показників або перший рівень для базисних показників.

Темп зміни показує інтенсивність зміни рівня ряду динаміки у відсотках або у разях (1 раз = 100 %).

Абсолютне значення 1 % відхилення ($A_{1\%}$) – це вагомість одного відсотка відхилення, яке визначається за формулою:

$$A_{1\%} = \frac{\Delta_a}{\Delta_{\%}} , \quad (9.5)$$

де Δ_a - абсолютне відхилення;

$\Delta_{\%}$ - відносне відхилення у відсотках.

Приклад розрахунку усіх перелічених показників, які характеризують ряди динаміки, наведено у табл. 9.1.

Таблиця 9.1 – Основні показники ряду динаміки

Показники	Роки			
	2006	2007	2008	2009
1	2	3	4	5
Реалізована продукція підприємства, млн. грн., (рівні ряду)	220	253	439	578
Абсолютні відхилення, млн. грн.				
- ланцюгові	X	33	186	139
- базисні	X	33	219	358
Відносні відхилення, %				
- ланцюгові	X	15,0	73,5	31,7
- базисні	X	15,0	99,5	162,7

Продовження табл. 9.1

1	2	3	4	5
Темпи змін, %				
- ланцюгові	100	115,0	173,5	131,7
- базисні	100	115,0	199,5	262,7
Абсолютні значення одного відсотка відносних відхилень, млн. грн.				
- ланцюгові	X	2,2	2,5	4,4
- базисні	X	2,2	2,2	2,2

Між ланцюговими і базисними темпами змін існує мультиплікативний взаємозв'язок – базисний темп змін дорівнює добутку ланцюгових темпів:

$$T_1 \times T_2 \times \dots \times T_n = \frac{y_2}{y_1} \times \frac{y_3}{y_2} \times \dots \times \frac{y_n}{y_{n-1}} = \frac{y_n}{y_1} = T_b, \quad (9.6)$$

де T_1, T_2, \dots, T_n – ланцюгові темпи змін;

T_b – базисний темп змін.

4. Вибір бази порівняння та середні темпи рядів динаміки

Аналізуючи ряди динаміки будь-яких статистичних показників, важливо **правильно вибрати базу порівняння**. Як видно з табл. 9.1, ланцюгові темпи зростання обсягів реалізованої продукції (115,0 %, 173,5 % і 131,7 %), а також відповідні ланцюгові відносні прирости (15,0 %, 73,5 % і 31,7 %), визначені по відношенню до різних баз порівняння (220 млн. грн., 253 млн. грн. та 439 млн. грн.). В результаті такого порівняння вагомість кожного відсотка (1 %) цих ланцюгових показників різко відрізняється одне від одного. Якщо за кожним відсотком зростання обсягу реалізованої продукції у 2007 р. порівняно з 2006 р. налічується тільки 2,2 млн. грн., то у 2009 р. у порівнянні з 2008 р. – у два рази більше: 4,4 млн. грн. І, навпаки, як видно з табл. 9.1, усі значення 1 % для відповідних базисних показників – однакові (2,2 = 2,2 млн. грн.).

Таким чином, для визначення дійсності динаміки статистичних показників у відсотках треба зіставляти не ланцюгові, а базисні відсотки. Тільки базисні відсотки порівнянні між собою і мають рівну вагомість 1 %. **Для бази порівняння треба вибирати лише кращі статичні показники,** враховуючи конкретні умови того чи іншого часу та території.

До кращих показників відносять міжнародні та національні стандарти, показники високо розвинутих країн, регіонів, міст, підприємств, організацій та установ. Не має ніякого сенсу вибирати недосконалу базу порівняння зі всіма її недоліками та невикористаними резервами.

Наприклад, прибуток двох аналогічних за розмірами і ресурсами фірм склав:

1) фірми «А» у поточному році – 1200 тис. грн., у минулому році – 1000 тис. грн.; темп зростання $\frac{1200}{1000} * 100 \% = 120\%$, приріст – 20 %;

2) фірми «Б» у поточному році – 200 тис. грн., у минулому – 100 тис. грн.; темп зростання $\frac{200}{100} * 100 \% = 200 \%$ (або 2 рази), приріст – 100 %.

На перший погляд, персонал фірми «Б» працював краще, ніж фірми «А», оскільки прибуток цієї фірми зріс у 2 рази (темп росту 200 %, приріст 100 %) проти 20 % у фірми «А». Але цей висновок помилковий. Відсотки зростання прибутку фірми «Б» і «А» порівнювати не можна, оскільки вони мають різну базу порівняння. Якщо вагомість 1 % прибутку фірми «А» складає 10 тис. грн. ($\frac{200}{20}$), то фірми «Б» – лише 1 тис. грн. ($\frac{100}{100}$). Для

коректного зіставлення прибутку цих двох фірм треба вибирати кращу базу порівняння. Краща база порівняння у минулому році була у фірми «А» – 1000 тис. грн. прибутку проти 100 тис. грн. у фірми «Б». Порівняно з минулим роком прибуток фірми «Б» був на цілих 80 % менший, ніж фірми «А» ($\frac{200}{1000} * 100 - 100 = -80\%$). Ці базисні відсотки – зменшення на

80 % і приріст на 20 % – мають однакову базу порівняння – 1000 тис. грн. Тому їх можна порівняти між собою. У даному випадку персонал фірми «А» працював краще, ніж фірми «Б».

Таким чином, порівнювати будь-які відсотки статистичних показників без попередньої обробки не можна. Їх треба зіставляти з однаковою базою порівняння.

Для узагальнення інтенсивності змін рядів динаміки розраховують **середній темп змін** за формулою середньої геометричної. Приклад застосування цієї формули наведено вище. Враховуючи мультиплікативний взаємозв'язок між ланцюговими і базисними темпами (9.6), в формулі середньої геометричної можна замінити добуток ланцюгових темпів

$(\prod_{i=1}^m T_i)$ базисним темпом (T_i) . У результаті отримаємо спрощену формулу середньої геометричної:

$$\bar{x} = \sqrt[n-1]{T_0} = \sqrt[m-1]{\frac{y_n}{y_1}}, \quad (9.7)$$

де \bar{x} – середній темп змін ряду динаміки;

T_0 – базисний темп змін;

y_n і y_1 – останній і перший рівні ряду динаміки;

n – число рівнів ряду динаміки.

За даними табл. 9.1 середній темп зростання обсягу реалізованої продукції підприємства становить:

$$\bar{x} = \sqrt[4-1]{\frac{578}{220}} = \sqrt[3]{2.627} = 1,380 \text{ або } 138,0\%.$$

Ряди динаміки, як правило, подаються не тільки в таблицях, але й показуються на графіках. При графічному зображенні динамічного ряду на осі абсцис відкладається шкала часу, а на осі ординат – шкала рівнів ряду.

5. Середні показники динаміки

Ряди динаміки складаються з багатьох варіаційних рівнів, а тому, як і будь-яка статистична сукупність, вони потребують деяких узагальнюючих характеристик. Для цього вираховують середні показники: середні рівні ряду, середні абсолютні зміни, середні темпи зміни і змінювання.

В інтервальному ряду з рівними інтервалами середній рівень ряду вираховується за формулою **середньої арифметичної простої**:

$$\bar{y} = \frac{\sum y_i}{n} = \frac{y_1 + y_2 + \dots + y_n}{n}, \quad (9.8)$$

де \bar{y} – середній рівень ряду;

$\sum y_i$ – сума рівнів ряду;

n – число рівнів.

Якщо окремі періоди інтервального ряду динаміки мають різну довжину, то для визначення середнього рівня використовують **середню арифметичну зважену**:

$$\bar{y} = \frac{\sum y_i t_i}{\sum t_i} = \frac{y_1 t_1 + y_2 t_2 + \dots + y_n t_n}{t_1 + t_2 + \dots + t_n}, \quad (9.9)$$

де y – рівні ряду;

t – проміжки часу.

Для визначення середнього рівня в моментному динамічному ряду з рівними інтервалами між сусідніми датами застосовують формулу **середньої хронологічної**.

$$\bar{y}_{xp.} = \frac{\frac{1}{2} y_1 + y_2 + y_3 + \dots + \frac{1}{2} y_n}{n-1}, \quad (9.10)$$

Середня абсолютна зміна визначається як середня арифметична проста з ланцюгових абсолютних змін за певні періоди і показує на скільки одиниць в середньому змінився рівень у порівнянні з попереднім.

$$\bar{\Delta}_y = \frac{\sum \Delta_y^L}{n}, \quad (9.11)$$

де $\bar{\Delta}_y$ – середня абсолютна зміна;

n – кількість змін.

Оскільки сума ланцюгових змін ($\sum \Delta_y^L$) дорівнює зміні за весь досліджуваний період ($y_n - y_1$), то формула середньої абсолютної зміни може мати наступний вигляд:

$$\bar{\Delta}_y = \frac{y_n - y_1}{n - 1}, \quad (9.12)$$

де y_n – кінцевий рівень ряду;

y_1 – початковий рівень ряду;

n – кількість рівнів динамічного ряду.

Середній темп змін розраховують за формулою **середньої геометричної**:

$$\bar{T}_p = \sqrt[n]{\prod(T_{pi})} = \sqrt[n]{T_{p1}^L \cdot T_{p2}^L \cdot T_{p3}^L \cdot \dots \cdot T_{pn}^L}, \quad (9.13)$$

де \bar{T}_p – середній темп змін;

$T_{p1}^L, T_{p2}^L, T_{p3}^L, \dots, T_{pn}^L$ – ланцюгові темпи змін;

n – число темпів.

Для обчислення середнього темпу змін використовують також іншу формулу:

$$\bar{T}_p = \sqrt[n-1]{\frac{y_n}{y_1}}, \quad (9.14)$$

де y_n – кінцевий рівень ряду;

y_1 – початковий рівень ряду;

n – кількість рівнів динамічного ряду.

Середні темпи змін обчислюють також за формулою **середньої геометричної зваженої**:

$$\overline{T}_p = \sqrt[t]{T_{p1}^{t_1} * T_{p2}^{t_2} * T_{p3}^{t_3} * ... * T_{pn}^{t_n}}, \quad (9.15)$$

де t – інтервал часу, на протязі якого зберігається даний темп змін;

Σt – сума відрізків часу періоду.

Середній темп змінювання (приросту) визначається як різниця між середнім темпом змін і одиницею (якщо середній темп змін у вигляді коефіцієнта), або 100 (якщо він у відсотках):

$$\overline{T}_{np} = \overline{T}_p - 1 \text{ (у вигляді коефіцієнтів)} \quad (9.16)$$

$$\overline{T}_{np} = \overline{T}_p - 100 \text{ (у вигляді відсотків)} \quad (9.17)$$

Застосування наведених показників динаміки є лише першим етапом аналізу рядів динаміки, який дозволяє виявити швидкість і інтенсивність розвитку явищ.

Подальший аналіз рядів динаміки соціально-економічних показників пов'язаний з більш складними узагальненнями, з визначенням основної тенденції, вивченням сезонних коливань рівнів і дослідженням зв'язку між рядами.



Питання для самоконтролю по темі 9

1. *Поняття про ряди динаміки.*
2. *Види рядів динаміки.*
3. *Дайте визначення понять „ряд динаміки”, „інтервальний ряд динаміки”, „моментний ряд динаміки”.*
4. *Елементи рядів динаміки.*
5. *Правила побудови рядів динаміки.*
6. *Основні характеристики рядів динаміки.*
7. *Перелічіть основні показники, які характеризують ряд динаміки.*
8. *Як визначається середній рівень інтервального і моментного ряду динаміки?*
9. *Визначення ланцюгових і базисних показників динаміки.*
10. *Чим відрізняються базисні показники ряду динаміки від ланцюгових?*
11. *За якими формулами треба визначати абсолютне та відносне відхилення рівнів ряду динаміки, а також їх темп зміни?*
12. *Вибір бази порівняння в рядах динаміки.*
13. *Який взаємозв'язок існує між ланцюговими і базисними темпами зміни?*
14. *Що показує абсолютне значення 1 % відхилення рівня ряду динаміки?*
15. *Як розраховують середній темп зміни рівнів ряду динаміки?*
16. *Охарактеризуйте методику вибору бази порівняння для правильної оцінки динаміки статистичних показників у відсотках.*

ТЕМА 10. АНАЛІЗ ТЕНДЕНЦІЙ РОЗВИТКУ

Питання для теоретичної підготовки

1. Визначення тенденції розвитку та прогнозування.
2. Вимірювання сезонних коливань.
3. Особливості вимірювання взаємозв'язків в рядах динаміки.

1. Визначення тенденції розвитку та прогнозування

Аналіз рядів динаміки є підставою та основним джерелом прогнозування розвитку економічних та соціальних явищ і процесів.



Статистичне прогнозування – це оцінка майбутнього стану явища на підставі глибокого статистичного аналізу тенденцій розвитку явищ, процесів та їх взаємозв'язків.

Виявлення основної тенденції (тренду) ряду, є одним з головних методів аналізу і узагальнення динамічних рядів. Для цього найчастіше використовують наступні способи:

- 1) спосіб збільшення інтервалів ряду динаміки;
- 2) спосіб ковзних середніх величин;
- 3) аналітичне вирівнювання .

Одним з найпростіших способів обробки ряду з метою виявлення закономірності зміни його рівнів є **спосіб збільшення інтервалів** (періодів) часу. Суть цього методу полягає у збільшенні інтервалів часу, які охоплюють рівні ряду динаміки. Для збільшених інтервалів часу дані динамічного ряду об'єднуються в групи по періодах і розраховується середній показник на період – триріччя, п'ятиріччя тощо. Ці рівні визначають основну тенденцію ряду динаміки.

Наприклад, є такі дані про виробництво продукції на підприємстві за десять років: 2000 – 260 млн. грн.; 2001 – 282; 2002 – 304; 2003 – 262; 2004 – 302; 2005 – 270; 2006 – 297; 2007 – 314; 2008 – 317; 2009 – 352 млн. грн.

Збільшимо інтервали часу з одного року до п'яти. В результаті отримаємо такі середні величини виробництва продукції на підприємстві:

За перші п'ять років (2000 – 2004 рр.):

$$\bar{x}_{2000-2004} = \frac{260 + 282 + 304 + 262 + 302}{5} = 282 \text{ млн. грн.};$$

За наступні п'ять років (2005 – 2009 рр.):

$$\bar{x}_{2005-2009} = \frac{270 + 297 + 314 + 317 + 352}{5} = 310 \text{ млн. грн.}$$

Збільшені середні демонструють основну тенденцію ряду динаміки – виробництво продукції на підприємстві зростає.

Важливим способом виявлення загальної тенденції ряду динаміки є **згладжування за допомогою ковзних середніх величин**. Тут також вдаються до укрупнення періодів, але воно проводиться шляхом послідовних зміщень на одну дату при збереженні постійного інтервалу періоду, тобто кожний наступний інтервал утворюється з попереднього послідовним зрушенням на один рівень.



Ковзні середні – це середні рівних збільшених інтервалів, які формуються шляхом заміни першого рівня інтервалу черговим рівнем ряду динаміки.

Таким чином, зсовуючись на один рівень, ковзні середні ніби ковзають по ряду динаміки від його початку до кінця. Так, за даними попереднього прикладу ковзні середні виробництва продукції становлять:

1) за перші п'ять років (2000-2004 рр.)

$$\bar{x}_{2000/04} = (260 + 282 + 304 + 262 + 302)/5 = 282 \text{ млн. грн.};$$

2) за 2001 – 2005 рр.:

$$\bar{x}_{2001/05} = (282 + 304 + 262 + 302 + 270)/5 = 284 \text{ млн. грн.};$$

3) 2002 – 2006 рр.:

$$\bar{x}_{2002/06} = (304 + 262 + 302 + 270 + 297)/5 = 287 \text{ млн. грн.};$$

4) 2003 – 2007 рр.:

$$\bar{x}_{2003/07} = (262 + 302 + 270 + 297 + 314)/5 = 289 \text{ млн. грн.};$$

5) 2004 – 2008 рр.:

$$\bar{x}_{2004/08} = (302 + 270 + 297 + 314 + 317)/5 = 300 \text{ млн. грн.};$$

6) 2005 – 2009 рр.:

$$\bar{x}_{2005/09} = (270 + 297 + 314 + 317 + 352)/5 = 310 \text{ млн. грн.}$$

Знайдені середні відносимо до середини відповідних інтервалів часу: перша середня – до 2002 р., друга – до 2003 тощо. Ковзні середні визначають основну тенденцію ряду динаміки (рис. 10.1).

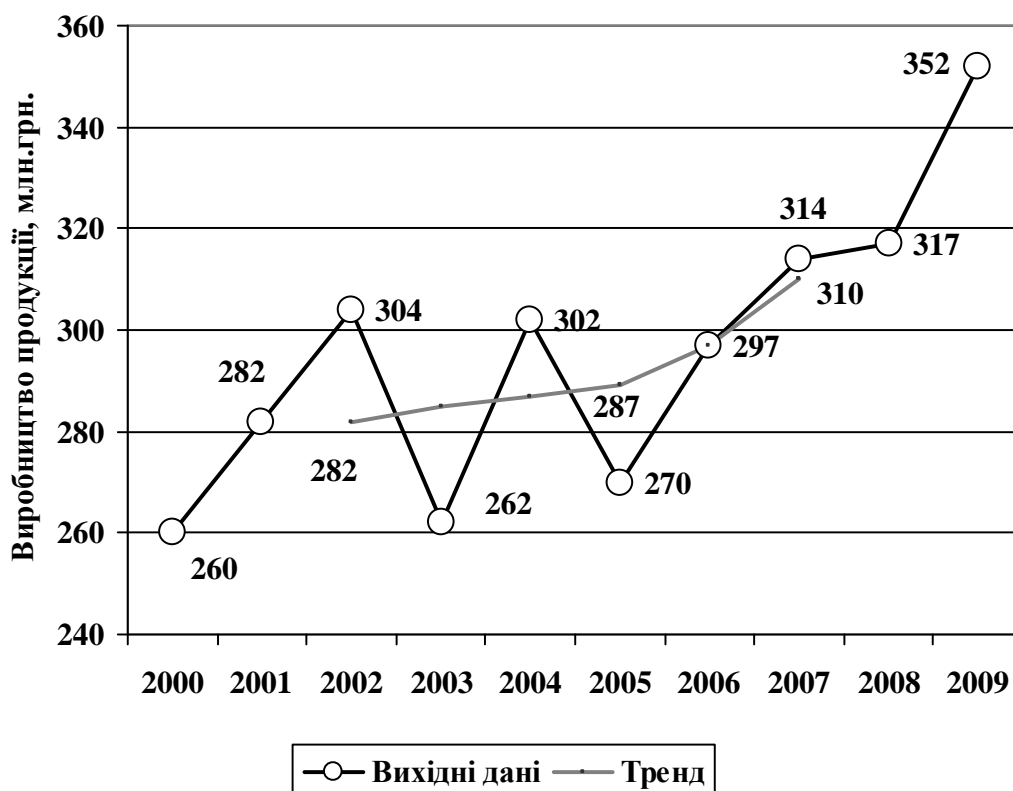


Рис. 10.1 – Динаміка виробництва продукції на підприємстві

На рис. 10.1 крива тренду визначена способом ковзних середніх. Вона показує основну тенденцію ряду динаміки – виробництво продукції на підприємстві зростає.

Спосіб ковзних середніх, на відміну від аналітичного вирівнювання, не потребує значного обсягу вихідної інформації (чималої кількості рівнів динаміки тощо).

Найбільш ефективним способом виявлення основної тенденції є аналітичне вирівнювання. **Аналітичне вирівнювання** – це визначення основної тенденції ряду динаміки на підставі всебічно обґрунтованого аналітичного рівняння.

Основні етапи аналітичного вирівнювання такі:

1. Виявлення особливостей **форми графіка** зміни рівнів ряду динаміки на підставі спільного, комплексного використання якісного (змістовного) й кількісного (числового) аналізу.
2. Вибір **форми аналітичного рівняння** за допомогою побудованого раніше попереднього графіка зміни рівнів ряду динаміки.
3. Обчислення параметрів аналітичного рівняння способом найменших квадратів.
4. Визначення прогнозних або планових величин рівнів ряду динаміки.
5. Оптимізація аналітичного рівняння за фактичними даними здійснення відповідних прогнозів і планів.

На практиці найбільш поширеними формулами, які виражають тенденцію розвитку (тренд) явищ, є: пряма, гіпербола, парабола другого порядку, показникова функція, ряди Фур'є, логістична функція, експонента та інші.

При виборі форми рівняння доцільно використовувати найбільш прості, а, отже, надійні лінійні рівняння, параболи 2-го ступеня, експоненти.

Вирівнювання за прямою використовується в тих випадках, коли абсолютні прирости більш-менш постійні, тобто коли рівні динамічного ряду змінюються в арифметичній прогресії, або близькі до неї.

$$y_t = a_0 + a_1 t, \quad (10.1)$$

де y_t – вирівняні рівні ряду динаміки (включаючи прогнози, які потрібно розрахувати; ці рівні залежать від кількості інтервалів часу (від числа років, кварталів тощо);

t – час (порядковий номер інтервалу або моменту часу);

a_0 і a_1 – параметри шуканої прямої (початковий рівень і щорічний приріст);

Наприклад, прибуток фірми за перший рік (2007 р.) становив 1,5 млн. грн., другий (2008 р.) – 3,5 млн. грн. та за третій (2009 р.) – 4 млн. грн. Попередній графік показує, що прибуток фірми характеризується стабільним абсолютним приростом рівнів. Тому припустимо, що прибуток змінюється за рівнянням прямої (10.1).

Параметри a_0 і a_1 прямої визначають способом найменших квадратів, згідно з яким сума квадратів відхилень вирівняних рівнів y_t від фактичних y_ϕ повинна бути мінімальною:

$$\sum_{t=1}^n (y_\phi - y_t)^2 = \min, \quad (10.2)$$

Враховуючи цю вимогу, параметри a_0 і a_1 знаходять за допомогою системи нормальних рівнянь для такої прямої лінії:

$$\begin{cases} na_0 + a_1 \sum t = \sum y_\phi \\ a_0 \sum t + a_1 \sum t^2 = \sum y_\phi t \end{cases}, \quad (10.3)$$

де y_ϕ – фактичні рівні ряду динаміки;

n – число рівнів.

Роки (t) послідовно позначаються як 1,2,3. Тобто $n = 3$; $\sum t = 6$; $\sum t^2 = 1 + 4 + 9 = 14$. У даному випадку:

$$\sum y_{\phi} = 1,5 + 3,5 + 4 = 9;$$

$$\sum y_{\phi} t = (1,5 \cdot 1) + (3,5 \cdot 2) + (4 \cdot 3) = 20,5.$$

Підставляючи фактичні дані наведеного прикладу в систему нормальних рівнянь, отримаємо:

$$\begin{cases} 3a_0 + 6a_1 = 9 \\ 6a_0 + 14a_1 = 20,5 \end{cases}$$

Щоб розв'язати цю систему, треба помножити перше рівняння на 2 і відняти його від другого: $2a_1 = 2,5$. Звідси $a_1 = 2,5 / 2 = 1,25$. Тоді $a_0 = 0,5$.
Шукане рівняння тренду:

$$Y_t = 0,5 + 1,25t.$$

Підставляючи в це рівняння порядкові інтервали часу, тобто роки (t), знайдемо вирівняні рівні ряду динаміки y_t : перший рік (2007 р.) – $y_{2007} = 0,5 + 1,25 \times 1 = 1,75$ млн. грн. прибутку; другий (2008р.) – 3 млн. грн.; третій (2009 р.) – 4,25 млн. грн.

Для прогнозного четвертого року (2010 р.) $t = 4$. Отже за прогнозом на майбутній рік прибуток фірми становитиме 5,5 млн. грн. ($0,5 + 1,25 \cdot 4$).

На більший термін, ніж один рік, розробляти прогноз за дуже спрощеним рівнянням прямої лінії немає сенсу. Таке складане явище, як прибуток, який залежить від величезного числа факторів, може охопити лише баланс підприємства, а не «наївне» рівняння прямої лінії.

Якщо зміни абсолютних приростів рівномірно збільшуються, то при виборі форми аналітичного рівняння використовують **параболу 2-го порядку**:

$$y_t = a_0 + a_1 t + a_2 t^2, \quad (10.4)$$

Параметри рівняння параболи другого порядку потрібно інтерпретувати наступним чином: « a_0 » – це величина, яка виражає середні умови утворення рівнів ряду; « a_1 » – швидкість розвитку даних динамічного ряду; « a_2 » – характеризує прискорення цього розвитку.

Вирівнювання за показниковою функцією проводиться в тих випадках, коли динамічний ряд розвивається в геометричній прогресії, тобто тоді, коли ланцюгові темпи росту більш-менш постійні.

При відносно стабільних темпах зростання застосовуються показову функцію:

$$y_t = a_0 a_1^t, \quad (10.5)$$

Коефіцієнт « a_1 », в показниковій функції характеризує середній темп зміни досліджуваної ознаки.

Основним методом аналітичного вирівнювання рядів динаміки слід вважати **кореляційно-регресійний метод**. Проте на практиці для прогнозування і планування рядів динаміки використовують головним чином балансовий і індексний метод.

Зокрема, усі найважливіші економічні і соціальні показники країн, регіонів, міст, підприємств та організацій всебічно обґрунтовують та визначають за допомогою відповідних прогнозних та планових балансів: бюджетів країн, регіонів, міст, бухгалтерських балансів підприємств тощо. Тому для вирішення економічних і соціальних проблем треба використовувати не якісь „ідеальні” моделі і методи, а цілеспрямовані інтегральні системи, в яких моделі і методи взаємодоповнюють один одне.

Вирівнювання рядів динаміки використовують також для знаходження відсутніх членів ряду за допомогою інтерполяції і екстраполяції.



Інтерполяцією називається в статистиці знаходження відсутнього показника усередині ряду.



Екстраполяцією в статистиці називається знаходження невідомих рівнів в кінці або на початку динамічного ряду.

Як інтерполяція, так і екстраполяція ґрунтуються на припущенні, що наявні величини цілком достатньо визначають темп розвитку досліджуваного явища і, отже, його можна поширювати на відсутні рівні динамічного ряду.

Коефіцієнт випередження являє собою відношення більшого середньорічного темпу приросту до меншого.

2. Вимірювання сезонних коливань



Сезонними коливаннями називаються більш-менш стійкі внутрішньорічні коливання в рядах динаміки, обумовлені специфічними умовами виробництва чи споживання певного виду продукції.

Для дослідження внутрішньорічних коливань можна використати цілий ряд методів (простої середньої, Персонса, рухомої середньої, аналітичного вирівнювання, рядів Фур'є), що забезпечують їх оцінку з різною точністю, надійністю і трудомісткістю.

Сезонні коливання характеризуються спеціальним показником, який називається індексом сезонності (I_s). У сукупності ці індекси утворюють сезону хвилю.



Індекс сезонності – це процентне відношення однойменних місячних (квартальних) фактичних рівнів динамічних рядів до їх середньорічних або вирівняних рівнів.

Індекс сезонності за методом простої середньої визначають за формулою

$$I_s = \frac{\bar{y}_i}{\bar{y}_z} * 100, \quad (10.6)$$

де I_s – індекс сезонності;
 y_i – середні місячні або квартальні рівні;
 y_z – загальна середня (місячна або квартальна).

Більшість динамічних рядів досліджуваних явищ мають тенденцію росту, тому для більш точного визначення сезонної хвилі в таких рядах необхідна нейтралізація еволюції тренду. З цією метою використовують **метод ланцюгових індексів (метод Персонса)**.

Метод Персонса дозволяє визначити зсув сезонної хвилі під впливом загальної тенденції розвитку. Цей зсув визначається як добуток середнього ланцюгового індексу за перший квартал на значення базисного індексу за четвертий квартал.

Для вивчення сезонності часто доводиться вираховувати **ковзну середню** з парним числом членів ряду, тому що характер динамічного ряду визначає тривалість періоду ковзної середньої, який повинен співпадати з періодом коливання, або бути кратним йому.

Згладжування за парним числом членів ряду незручне тим, що середня мусить бути віднесена тільки до середини між двома датами, тобто проходить зсув періоду, до якого відноситься рівень. Усунення зсуву періоду проводять способами перетворення рівнів і центруванням.

Для розрахунку сезонних коливань використовують також метод аналітичного вирівнювання за рівнянням прямої.

Параметри « a_0 » і « a_1 » для рівняння прямої $y_t = a_0 + a_1 t$ знаходяться методом найменших квадратів розв'язуючи систему нормальних рівнянь.

3. Особливості вимірювання взаємозв'язків в рядах динаміки

При вивченні кореляційних зв'язків в багатомірних рядах динаміки спостерігається певна залежність рівнів даного періоду від попереднього, внаслідок чого виникають певні методологічні особливості. В таких динамічних рядах фактором зміни рівнів виступає, крім інших, також час.

Вплив даного рівня динамічного ряду на зміну наступного з плином часу приводить до **автокореляції**.

У практиці економіко-статистичного аналізу рядів динаміки застосовують різні способи усунення автокореляції, такі як спосіб різницевих перетворень (при лінійному тренді), спосіб відхилень тенденції (при нелінійній залежності), або введення змінної величини « t » у рівняння регресії $y_t = f(x_1, x_2, x_3, \dots, t)$, де вона відіграє роль чинника часу.

При застосуванні методу **регресії** для дослідження динамічних рядів виникає особливість, яка заключається в тому, що в рівнях динамічних рядів присутня **авторегресія**, яка проявляється так же, як і автокореляція.

Авторегресія виражає залежність величини рівня динамічного ряду від попередніх значень рівня в певні моменти часу. В цьому випадку рівняння регресії повинно виключити авторегресію.

Коефіцієнт автокореляції відхилень набуває значення в межах від -1 до +1 і визначається за формулою

$$r_a = \frac{\sum \varepsilon_t \cdot \varepsilon_{t+1}}{\sum \varepsilon_t^2}, \quad (10.7)$$

де $\varepsilon_t = y - y_t$;

y – фактичні дані;

y_t – вирівняні дані.

У багатьох економіко-статистичних дослідженнях доводиться вивчати паралельно декілька динамічних рядів, в яких коливання рівнів взаємообумовлені, наприклад, динаміка цін на електроенергію на оптовому ринку в значній мірі пов'язана з цінами на паливо; у свою чергу, динаміка цін на

паливо залежить від динаміки їх видобутку, обсягів їх експорту; попит населення на певні товари народного споживання залежить від пропозиції, тобто від обсягів їх виробництва тощо.

Для вимірювання залежності між такими рядами динаміки використовують методи кореляції, тобто розраховують різні коефіцієнти кореляції.

Одним із способів усунення автокореляції є **корелювання відхилень фактичних рівнів від вирівняних**, які відображають тренд. Для цього потрібно:

- а) провести аналітичне вирівнювання порівнювальних рядів;
- б) визначити величини відхилень кожного фактичного рівні динаміки від їх вирівняних значень;
- в) провести корелювання отриманих відхилень.

Коефіцієнт кореляції відхилень розраховують за формулою:

$$r = \frac{\sum d_x \cdot d_y}{\sqrt{\sum d_x^2 * \sum d_y^2}}, \quad (10.8)$$

де r – коефіцієнт кореляції відхилень фактичних рядів від вирівняних;

$d_x = x - x_t$; $d_y = y - y_t$ – відхилення фактичних рівнів динаміки від вирівняних обох рядів.

До аналогічних результатів можна прийти, якщо знайти **кореляцію різниць між наступними і попередніми рівнями** обох рядів. При заміні рівнів динамічних рядів різницями між ними усувається вплив автокореляції у кожному динамічному ряду.

Коефіцієнт кореляції перших різниць визначається за формулою:

$$r = \frac{\sum \Delta x * \Delta y}{\sqrt{\sum \Delta x^2 * \sum \Delta y^2}}, \quad (10.9)$$

де r – коефіцієнт кореляції перших різниць;

$\Delta x = x_i - x_{i-1}$; $\Delta y = y_i - y_{i-1}$ – різниці між наступними і попередніми рівнями обох рядів.



Питання для самоконтролю по темі 10

1. *Визначіть поняття „статистичне прогнозування”.*
2. *Визначіть поняття „тенденція (тренд) ряду”.*
3. *Виявлення закономірностей зміни рівнів ряду.*
4. *Спосіб збільшення інтервалів ряду динаміки.*
5. *Спосіб ковзних середніх величин.*
6. *Аналітичне вирівнювання.*
7. *Основні етапи аналітичного вирівнювання.*
8. *Основні форми рівнянь аналітичного вирівнювання.*
9. *Визначіть поняття „інтерполяція” і „екстраполяція”.*
10. *Визначіть поняття „сезонні коливання”.*
11. *Визначіть поняття „індекс сезонності”.*
12. *Метод Персона.*
13. *Особливості взаємозв'язків в рядах динаміки.*
14. *Визначіть поняття „авторегресія” і „автокореляція”.*

ТЕМА 11. ІНДЕКСИ

Питання для теоретичної підготовки

1. Поняття індексів. Основні елементи індексів.
2. Індивідуальні й загальні індекси.
3. Агрегатна форма загального індексу.
4. Середні індекси.
5. Інденси з постійними і змінними вагами.
6. Взаємозв'язки індексів.
7. Територіальні індекси.

1. Поняття індексів. Основні елементи індексів

В економічній науці ще не знайдена (і, можливо, ніколи не буде знайдена) єдина міра, еталон, що був би застосовний до різних економічних величин і дозволяв би вимірювати їх – це головна особливість економічних величин взагалі.

У господарській практиці вимір економічних величин знаходить свій вираз у **статистичних показниках**, що характеризують властивості економічних об'єктів, і які необхідні для управління та наукового пізнання.

Найпоширенішим у практиці статистичної роботи поряд з методом середніх величин слід визнати **індексний метод**, тобто сукупність прийомів для виявлення загальних показників і виміру динаміки у складних статистичних сукупностях.

При цьому під **складною (різнозіставною)** розуміється така статистична сукупність, окремі елементи якої мають різну натурально-речову форму.

Отримані на підставі цього методу показники використовуються для характеристики розвитку аналізуємих показників у часі, по території,

вивчення структури й взаємозв'язків, виявлення ролі факторів у зміні складних явищ.



Індекс (від лат. *index* – покажчик, список) – статистичний показник, що характеризує у відносному вигляді кількісну зміну економічних величин, параметрів економічних і соціальних процесів за певний період часу.

Індекси – це відносні системні показники, які характеризують зміни економічних, соціальних та інших явищ у часі, у просторі чи у зіставленні з будь-якою базою порівняння (стандартними, плановими або середніми величинами, показниками минулих періодів, кращих підприємств, організацій, установ тощо).

Як правило, статистичні показники, що складаються з різнорідних елементів, безпосереднє підсумовування яких неможливо через їх несумісності. Тому в *основу індексного методу* при визначенні змін у досліджуваному явищі покладений перехід від натурально-речової форми вираження до вартісних вимірників (ціна, собівартість, трудомісткість тощо).

Сумісність і логіка порівняння окремих показників досягається за допомогою їх вартісного вираження.

Для обчислення будь-якого індексу, за визначенням, необхідно зробити зіставлення не менше двох величин. Основним елементом **індексного відношення** є *індексуєма величина*.



Під **індексуємою величиною** розуміється значення ознаки статистичної сукупності, зміна якої є об'єктом вивчення.

При обчисленні індексів розрізняють **порівнювану** величину (чисельник індексного відношення), і величину, з якою провадиться порівняння, – **базисну**. Вибір бази порівняння визначається *метою* дослідження.

Для індексів, що характеризують зміну індексуємої величини у *часі*, за базисну приймають розмір показника в періоді, що передує звітному.

При *територіальних* порівняннях за базу приймають дані по частині території.

Для індексів, що характеризують *виконання плану*, за базу приймають величину планового завдання.

Основними елементами кожного індексу є:

- а) індексуєма величина;
- б) тип (форма) індексу;
- в) ваги індексу;
- г) терміни обчислення.

Згідно з таким розподілом:

- а) залежно від ***індексуємої величини*** розрізняють: індекси цін, індекси фізичного (натурального) обсягу продукції, індекси продуктивності праці тощо;
- б) залежно від ***методології розрахунку (форми)*** розрізняють: індекси агрегатні й індекси середні, у тому числі, індекси середні арифметичні, індекси середні геометричні, індекси середні гармонійні тощо;
- в) залежно від ***ваги*** розрізняють: індекси прості (незважені) і індекси зважені, а серед останніх – індекси з постійними (незмінними) вагами й індекси зі змінними вагами;
- г) залежно від ***термінів обчислення*** розрізняють: індекси базисні (з постійною, незмінною у часі базою) і індекси ланцюгові (якщо числові значення індексуємої величини в кожний даний „поточний” термін зіставляються з їх значеннями у попередній термін, інакше, індекс зі змінною базою);

До **основних завдань**, які розв'язуються за допомогою індексів, слід віднести:

- 1) характеристика загальної зміни складного економічного показника або формуючих його окремих показників-факторів;

- 2) виділення у зміні складного показника впливу одного з факторів шляхом елімінування (виключення в процесі аналізу) впливу інших факторів.

2. Індивідуальні й загальні індекси

За ступенем охоплення підлягаючих узагальненню одиниць досліджуваної сукупності індекси підрозділяються на *індивідуальні (елементарні) й загальні (зведені)*.

Індивідуальні індекси характеризують зміни тільки окремих елементів статистичної сукупності (наприклад зміна видобутку вугілля марки АШ у загальному обсязі антрациту). Обов'язковою умовою для його обчислення є максимальна однорідність об'єкту, для якого він обчислюється. Індивідуальний індекс є відносною величиною динаміки, і отже може бути ланцюговим або базисним.

Індивідуальні індекси прийнято позначати символом «*i*».

Загальні індекси відображають зведену динаміку усіх елементів, що утворюють різнозіставну статистичну сукупність.

Статистична сукупність є різнозіставною за даною ознакою, якщо підсумкову величину цієї ознаки у всій сукупності прямим, безпосереднім підсумовуванням його значень окремих одиниць обчислити не можна (наприклад, натуральна величина продукції, що складається з речовинно різних фізичних одиниць або частин) або якщо таке підсумовування хоч і можливе формально, але приводить до результату, позбавленого економічного змісту (наприклад, сума цін речовинно різних товарів, взятих лише по одній одиниці натурального виміру).

Індекси можуть бути обчислені не тільки для всієї різнозіставної сукупності (загальні), але й для будь-якої її характерної частини, істотної групи одиниць (групові індекси або *субіндекси*), наприклад: загальний індекс опто-

вих цін усіх взагалі товарів та групові індекси цін товарів продовольчих і непродовольчих, промислових і сільськогосподарських тощо.

Загальні індекси прийнято позначати символом « I ».

При вивченні зміни цін індексуємою величиною є ціна одиниці товару « p ». При вивченні зміни фізичного обсягу товарної маси індексуємою величиною виступають дані про кількість товарів у натуральних вимірниках – « q ».

Індивідуальні індекси фізичного обсягу реалізації товарів i_q визначаються за формулою:

$$i_q = \frac{q_1}{q_0}, \quad (11.1)$$

де q_1 і q_0 – кількість продажів окремого товарного різновиду у поточному й базисному періодах у натуральних вимірниках.

Для визначення **індивідуальних індексів цін** i_p застосовується формула:

$$i_p = \frac{p_1}{p_0}, \quad (11.2)$$

де p_1 і p_0 – ціни за одиницю товару в поточному і базисному періодах.

Результат розрахунку індексних відношень може бути виражений коефіцієнтом, у відсотках або проміле.

Індекси створюють **системи індексів**, тому індекси є системними показниками. Найчастіше індекси створюють системи індексів-співмножників. Існує



Правило взаємозв'язку індексів: індекси пов'язані між собою так само, як і показники, з яких вони вираховуються.

Наприклад, добуток кількостей товарів (q_i) на їх ціни (p_i) дорівнює вартості цих товарів (Q). Так само пов'язані й індекси цих трьох показників:

$$i_q \times i_p = i_Q. \quad (11.3)$$

Загальний індекс фізичного обсягу реалізації товарів I_q визначається за формулою:

$$i_q = \frac{\sum q_i^1 p_i^0}{\sum q_i^0 p_i^0}, \quad (11.4)$$

де $\sum q_i^1 p_i^0$ – обсяг продажу (у вартісному виразі) товарів у поточному періоді за цінами базисного періоду;

$\sum q_i^0 p_i^0$ – обсяг продажу (у вартісному виразі) товарів у базисному періоді за цінами базисного періоду;

Загальний індекс цін I_p розраховується за наступною формулою:

$$i_p = \frac{\sum q_i^1 p_i^1}{\sum q_i^1 p_i^0}, \quad (11.5)$$

де $\sum q_i^1 p_i^1$ – вартість продажу товарів у поточному періоді за цінами того ж поточного періоду;

$\sum q_i^1 p_i^0$ – вартість продажу товарів у поточному періоді за цінами базисного періоду;

Загальний індекс вартості продажу товарів визначається як:

$$i_{qp} = \frac{\sum q_i^1 p_i^1}{\sum q_i^0 p_i^0}, \quad (11.6)$$

де $\sum q_i^1 p_i^1$ – вартість продажу товарів у поточному періоді за цінами того ж поточного періоду;

$\sum q_i^0 p_i^0$ – вартість продажу товарів в базисному періоді за цінами базисного періоду.

3. Агрегатна форма загального індексу

Основна складність порівняння показників різнозіставної сукупності полягає у відшуванні методу їхнього порівняння. Досягнення порівнянності різнорідних одиниць у таких сукупностях здійснюється введенням в індексні відношення спеціальних співмножників індексуємих величин, які називають *співвимірниками*. Вони необхідні для переходу від натуральних вимірників різнорідних одиниць статистичної сукупності до однорідних показників. При цьому в чисельнику й знаменнику загального індексу змінюється лише значення індексуємої величини, а їх співвимірники є постійними величинами й фіксуються на одному рівні (поточного або базисного) періоду. У цьому випадку на величині індексу позначається лише вплив фактору, що визначає зміну індексуємої величини.

У якості співвимірників індексуємих величин виступають тісно пов'язані з ними економічні показники: ціни, тарифи тощо. Добуток кожної індексуємої величини на співвимірник утворить в індексному відношенні певні економічні категорії (товарообіг, виторг від продажу тощо).

Основною формою загальних індексів є **агрегатні індекси** (від лат. „*aggrega*” – „приєдную”). У чисельнику й знаменнику загальних індексів в агрегатній формі містяться **з'єднані набори (агрегати)** елементів досліджуваних статистичних сукупностей.

При визначенні загального індексу цін в агрегатній формі I_p у якості співвимірника індексуємих величин p^I і p^0 можуть застосовуватися дані про обсяги реалізації товарів у поточному періоді q^I .

При множенні q^I на індексні величини:

- у чисельнику індексного відношення утвориться значення $\sum p^1_i q^1_i$, тобто вартість продажів обсягів товарів поточного періоду за цінами того ж поточного періоду;

- у знаменнику індексного відношення утвориться значення $\sum p_i^1 q_i^0$, тобто умовна вартість продажів обсягу товарів поточного періоду за цінами базисного періоду.

Агрегатна формула такого загального індексу має такий вигляд:

$$I_p = \frac{\sum p_i^1 q_i^1}{\sum p_i^0 q_i^1}, \quad (11.7)$$

Індекс Пааше (*Paasche price index*) – один з найбільш розповсюджених агрегатних індексних показників, за яким оцінюється зміна загального рівня фізичних обсягів або цін (наприклад роздрібних). Обчислюється з «вагами» звітного періоду (**формула Пааше**):

Фізичного обсягу:

$$I_q = \frac{\sum q_i^1 p_i^1}{\sum q_i^0 p_i^1}$$

Цін:

$$I_p = \frac{\sum p_i^1 q_i^1}{\sum p_i^0 q_i^1}, \quad (11.8)$$

При порівнянні чисельника й знаменника формул (11.7) у їх різниці визначається показник абсолютної зміни (приросту чи зменшення) відповідного показника.

Зміни кількості реалізованих товарів (у вартісному вимірі) (Δ_{pq}) за рахунок фактора зміни обсягів виробництва в поточному періоді в порівнянні з базисним періодом *в цінах поточного періоду*:

$$\Delta_{pq}(q) = \sum q_i^1 p_i^1 - \sum q_i^0 p_i^1, \quad (11.9)$$

Зміни товарного обігу (Δ_{qp}) за рахунок фактору зміни цін у поточному періоді в порівнянні з базисним періодом:

$$\Delta_{qp}(p) = \sum p_i^1 q_i^1 - \sum p_i^0 q_i^1, \quad (11.10)$$

При іншому способі визначення агрегатного індексу зміни загального рівня фізичних обсягів або цін, німецьким економістом Е.Ласпейре-

сом була запропонована формула з „вагами” звітнього періоду. Тому індекси, розраховані по цій формулі, прийнято називати **індексами Ласпейреса**.

Фізичного обсягу:

Цін:

$$Iq = \frac{\sum q_i^1 p_i^0}{\sum q_i^0 p_i^0} \quad I_p = \frac{\sum p_i^1 q_i^0}{\sum p_i^0 q_i^0}, \quad (11.11)$$

При порівнянні чисельника й знаменника формул (11.10) у їх різниці також визначається показник абсолютної зміни (приросту чи зменшення) відповідного показника.

Зміни кількості реалізованих товарів (у вартісному вимірі) (Δ_{pq}) за рахунок фактору зміни обсягів виробництва у поточному періоді в порівнянні з базисним періодом *в цінах базисного періоду*:

$$\Delta_{pq}(q) = \sum q_i^1 p_i^0 - \sum q_i^0 p_i^0, \quad (11.12)$$

Зміни товарного обігу (Δ_{qp}) за рахунок фактору зміни цін у поточному періоді в порівнянні з базисним періодом:

$$\Delta_{qp}(p) = \sum p_i^1 q_i^0 - \sum p_i^0 q_i^0, \quad (11.13)$$

Застосування індексів Пааше й Ласпейреса залежить від мети дослідження. При побудові формул агрегатних індексів використовують наступне **правило**:



Якщо індексована величина – якісний показник, який знаходять шляхом ділення (ціна, собівартість, продуктивність тощо), то беруться ваги звітнього періоду, а якщо індексована величина – кількісний показник, який можна підсумувати (фізичний обсяг продукції, чисельність працівників тощо), беруться ваги базисного періоду.

З метою усунення недоліків, притаманних індексам Пааше та Лайспереса розраховується їх середня геометрична величина – індекс Фішера:

$$I_F = \sqrt{I_P * I_L} , \quad (11.14)$$

При синтезуванні загального індексу цін замість фактичної кількості товарів (у звітному або базисному періодах) у якості співвимірника індексуємих величин (p_1 і p_0) можуть застосовуватися середні величини обсягів реалізації товарів (\bar{q}) за два або більше періодів. При такому способі розрахунку формула загального індексу синтезується у наступному вигляді:

$$I_p = \frac{\sum p_1 \bar{q}}{\sum p_0 \bar{q}} , \quad (11.15)$$

де \bar{q} – середній обсяг реалізації товарів за періоди, що аналізуються.

У літературі такий індекс прийнято називати **індексом Лоу**.

Якщо при визначенні цього індексу цін вихідна інформація містить лише дані про кількість реалізації товарів у базисному й поточному періодах, то середня їх величина визначається методом середньої незваженої:

$$\bar{q} = \frac{q_0 + q_1}{2} , \quad (11.16)$$

Індекс цін Лоу застосовується у розрахунках при аналізі закупівель або реалізації товарів, робіт і послуг протягом тривалих періодів часу (п'ятирічка тощо). Цей метод надає можливість здійснювати аналіз цін з урахуванням змін в асортиментному складі товарів, що відбуваються всередині окремих субперіодів.

Наведена методика визначення загальних індексів цін в агрегатній формі може бути застосована й до інших індексів якісних показників: собівартості I_z , продуктивності праці I_t тощо.

4. Середньозважені індекси

У деяких випадках загальні індекси обчислюють як середні арифметичні, перетворені з відповідних агрегатних індексів.

Перетворюють агрегатний індекс в середній з індивідуальних індексів, підставляючи у його чисельник або знаменник замість індексованого показника його вираз, виведений з формули індивідуального індексу. Якщо таку заміну роблять у чисельнику, то агрегатний індекс перетворюється на *середній арифметичний*, а якщо у знаменнику – *середній гармонічний*.

Перетворимо агрегатний індекс фізичного обсягу на середній арифметичний. Як відомо, формула загального індексу фізичного обсягу має

вигляд: $I_q = \frac{\sum q_i^1 p_i^0}{\sum q_i^0 p_i^0}$, а індивідуальний індекс фізичного обсягу

$$i_q = \frac{q_1}{q_0}, \text{ звідки } q_1 = i_q * q_0.$$

Замінивши в формулі агрегатного індексу фізичного обсягу продукції індексовану величину q_1 на $i_q * q_0$, отримаємо формулу **середнього арифметичного індексу** фізичного обсягу продукції:

$$\bar{I}_q = \frac{\sum i_q q_i^0 p_i^0}{\sum q_i^0 p_i^0}, \quad (11.17)$$

Даний індекс виступає як середня арифметична величина з індивідуальних індексів, зважених на вартість продукції базисного періоду ($q_0 * p_0$).

Щоб перетворити агрегатний індекс цін у середній гармонічний, потрібно в знаменнику агрегатного індексу цін замінити індексовану величину p_0 її вираженням із індивідуального індексу $p_0 = \frac{p_1}{i_p}$, звідки $p_0 = \frac{p_1}{i_p}$ – отримуємо

формулу середнього гармонічного індексу цін наступного вигляду:

$$\bar{I}_p = \frac{\sum p^1_i q^1_i}{\sum \frac{p^1_i q^1_i}{i_p}}, \quad (11.18)$$

Цей індекс являє собою середню гармонічну з величин, обернених індивідуальним індексам цін, зважених на вартість продукції у поточному періоді ($p_1 * q_1$).

Середні арифметичні і гармонічні індекси повинні співпадати за своєю величиною з відповідними агрегатними індексами.

Вибір форми індексу залежить від поставленого завдання дослідження і від наявності даних, необхідних для обчислення того чи іншого індексу.

5. Індекси з постійними й змінними вагами

При вивченні динаміки соціально-економічних явищ часто доводиться здійснювати індексні зіставлення більш ніж за два періоди. Тому індексні величини можуть визначатися як на постійній, так і на змінній базис порівняння. При цьому якщо завдання аналізу полягає в одержанні характеристик зміни досліджуваного явища у всіх наступних періодах у порівнянні з початковим, то обчислюються базисні індекси. Наприклад, зіставлення обсягу роздрібного товарообігу II, III і IV кварталів з I кварталом.

Але якщо потрібно охарактеризувати послідовність досліджуваного явища з періоду в період, то обчислюються ланцюгові індекси. Наприклад, при вивченні обсягу роздрібного товарообігу за кварталами року зіставляють товарообіг II кварталу з I кварталом, III кварталу – з II кварталом і IV кварталу – з III кварталом.

Залежно від завдання дослідження й характеру вихідної інформації обчислюють базисні й ланцюгові індекси як індивідуальні, так і загальні.

Способи розрахунку індивідуальних базисних і ланцюгових індексів аналогічні розрахунку відносних величин динаміки.

Загальні індекси залежно від їх виду (за економічним змістом) обчислюють зі змінними показниками. Так, розглянута у попередніх питаннях агрегатна форма загального індексу фізичного обсягу обчислюються як індекс з постійними вагами-співвимірниками. Агрегатна форма загального індексу цін обчислюється як індекс із змінними вагами-співвимірниками.

Ланцюгові й базисні індекси з постійними вагами-співвимірниками перебувають у наступному взаємозв'язку:

- 1) добуток ланцюгових індексів дає базисний індекс (останнього періоду):

$$I_1^{\text{л}} * I_2^{\text{л}} * \dots * I_{n-1}^{\text{л}} * I_n^{\text{л}} = \frac{i_2}{i_1} \times \frac{i_3}{i_2} \times \dots \times \frac{i_n}{i_{n-1}} = \frac{i_n}{i_1} = I_{\text{б}}, \quad (11.19)$$

- 2) відношення наступного базисного індексу до попереднього базисного індексу дає ланцюговий індекс (наступного періоду).

6. Взаємозв'язки індексів

Досліджувані в статистиці торгівлі показники перебувають між собою в певному зв'язку. Так, для кожного періоду обсяг роздрібного товарообігу залежить від кількості реалізованих товарів і від рівня цін на ці товари. Зв'язок між змінами обсягу товарообігу, кількістю продажу товарів і рівнем їх цін виражається в системі взаємозалежних індексів товарообігу.

Оскільки величина обсягу товарообігу дорівнює добутку кількості продажу товарів на ціни, то індекс фізичного обсягу I_q , помножений на індекс цін I_p , дає індекс товарообігу у фактичних цінах I_{qp} :

$$I_q * I_p = I_{qp}, \quad (11.20)$$

Значення формули полягає в тому, що на її основі виявляється вплив окремих факторів на зміну товарообігу.

На підставі цієї формули можна визначити зміну товарообігу в незмінних цінах:

$$I_q = I_{qp} / I_p, \quad (11.21)$$

На основі формули можна також по відомих індексах товарообігу у фактичних цінах I_{qp} і товарообігу у порівнянних цінах I_q визначити індекс цін I_p :

$$I_p = I_{qp} / I_q, \quad (11.22)$$

При використанні даних формул взаємозалежних індексів треба мати на увазі, що взаємозв'язок утворюється лише за умови, коли ваги-співвимірники в індексах фізичного обсягу й цін беруться на різних рівнях.

Індекс впливу структурних зрушень у реалізованій продукції на зміну середньої ціни визначається за формулою:

$$I_{\tilde{n}\partial\partial} = \frac{\sum p_0 q_1}{\sum q_1} : \frac{\sum p_0 q_0}{\sum q_0}, \quad (11.23)$$

Індекс цін постійного (фіксованого) складу розраховується:

$$I_p = \frac{\sum p_1 q_1}{\sum p_0 q_1}, \quad (11.24)$$

Обчислені за формулами індекси перебувають у взаємозв'язку:

$$\overline{I_p} = I_p * I_{\tilde{n}\partial\partial}, \quad (11.25)$$

7. Територіальні індекси

У сучасних умовах розвитку статистики все більшого значення здобуває використання індексного методу для територіальних порівнянь.

Загальні принципи використання індексного методу при територіальних порівняннях багато в чому подібні до вивчення динаміки складних статистичних сукупностей. Так, при двосторонніх порівняннях кожний регіон може бути прийнятий як у якості порівнюваного, та і як база порів-

няння. При цьому для визначення зведених (загальних) індексів необхідно вирішити питання про ваги-співвимірювачі індексуємих величин.

Для аналізу співвідношення рівнів цін на товари, реалізовані в місті К у порівнянні з містом М, визначається зведений (загальний) індекс цін, у якому як ваги-співвимірювачі індексуємих величин P_K і P_M приймаються кількості товарів, проданих у місті К:

$$I_{P_K/M} = \frac{\sum q_k P_k}{\sum q_k P_M}, \quad (11.26)$$

У формулі чисельник $\sum q_k P_k$ характеризує фактичний обсяг товарообігу при продажу даних асортиментів товарів у місті К (за сформованими там цінами). Знаменник формули $\sum q_k P_M$ відображає умовну величину товарообігу, що могла бути при продажу досліджуваних асортиментів товарів за цінами, що склалися у місті М.

Різниця між чисельником і знаменником формули відображає суму економічного ефекту від розходження цін у даних містах:

$$\sum q_k P_k - \sum q_k P_M, \quad (11.27)$$

Але можлива й інша постановка мети аналізу: визначити співвідношення рівнів цін на товари, реалізовані у місті М у порівнянні з містом К. Для визначення зведеного (загального) індексу цін як ваги-співвимірювачі індексуємих величин використовуються дані про кількість реалізації товарів у місті М (q_m):

$$I_{P_M/K} = \frac{\sum q_m P_M}{\sum q_m P_K}, \quad (11.28)$$

У формулі чисельник індексного відношення $\sum q_m P_M$ відображає фактичний обсяг товарообігу реалізації товарів у місті М (за сформованими там цінами), а знаменник індексного відношення $\sum q_m P_K$ характеризує умовну величину товарообігу, що міг би утворитися при продажу досліджуваних асортиментів товарів за цінами міста К.

Зіставленням у різниці чисельника й знаменника індексу визначається сума економічного ефекту від розходження у рівнях цін за даними регіонами:

$$\sum q_m p_m - \sum q_m p_k , \quad (11.29)$$

Для подолання суперечливих показань між зведеними (загальними) територіальними й індивідуальними (однотоварними) індексами визначається індекс цін, у якому як вага-співвимірник виступає сума реалізації товарів по двох регіонах (містах) q :

$$q = q_k + q_m , \quad (11.30)$$

З урахуванням даного значення формули зведеного (загального) індексу цін при аналізі зміна цін у місті К у порівнянні з містом М наступна:

$$I_{p_k/m} = \frac{\sum p_k q}{\sum p_m q} , \quad (11.31)$$

Зміни цін у місті М у порівнянні з містом К визначається за формулою:

$$I_{p_m/k} = \frac{\sum p_m q}{\sum p_k q} , \quad (11.32)$$

У вільні (загальних) територіальних індексах фізичного обсягу як ваги-співвимірники можуть виступати середні ціни \bar{p} :

В індексах зі змінними вагами-співвимірниками такої залежності немає. Так, добуток ланцюгових індексів не дає базисного індексу.



Питання для самоконтролю по темі 11

1. *Поняття про статистичні індекси.*
2. *Надайте визначення та приклади різнозіставних сукупностей.*
3. *Поняття про індексне відношення.*
4. *Основні елементи індексів.*
5. *Головні завдання, які вирішуються за допомогою індексів.*
6. *Поняття про індивідуальні індекси, приклади.*
7. *Поняття про загальні індекси, приклади.*
8. *Базисні й ланцюгові індекси.*
9. *Загальний індекс фізичного обсягу продукції.*
10. *Загальний індекс цін.*
11. *Поняття про співвимірювачи в індексах.*
12. *Поняття про агрегатну форму загального індексу.*
13. *Індекси Пааше.*
14. *Індекси Лайспереса.*
15. *Правило використання та вибору форм індексів.*
16. *Середньозважені індекси.*
17. *Індекси з постійними та змінними вагами.*
18. *Взаємозв'язок індексів.*
19. *Поняття про територіальні індекси.*

ТЕМА 12. ПОДАННЯ РЕЗУЛЬТАТІВ СТАТИСТИЧНИХ ДОСЛІДЖЕНЬ

Питання для теоретичної підготовки

1. Статистичні таблиці.
2. Поняття статистичного графіку. Роль графічного методу у статистиці.
3. Складові елементи статистичного графіка й правила його побудови.
4. Цілі й завдання, що вирішуються за допомогою графіків.
5. Форми й види графіків.

1. Статистичні таблиці

Результати статистичного зведення і групування обов'язково викладаються у формі статистичних таблиць. Однак у практичній статистичній роботі застосовування таблиць не обмежується тільки висвітленням цих результатів. Важливу роль вони відіграють у всій економічній роботі, тому складання статистичних таблиць слід вважати заключним етапом будь-якого статистичного дослідження, на підставі якого наочно можна зробити вірні висновки.

Найважливішою вимогою, яка пред'являється до статистичних таблиць, слід вважати подання результатів досліджуваного матеріалу у наглядній для отримувача формі.

Основною особливістю табличного викладення матеріалу є те, що показники, які характеризуються в статистичній таблиці, можна об'єднати під одним загальним заголовком.

Проте не всяку таблицю слід вважати статистичною. На відміну від допоміжних розрахункових таблиць (логарифмів, коефіцієнтів, множення тощо) статистичними таблицями можуть вважатися тільки ті, що містять

інформацію про результати статистичного аналізу соціально-економічних явищ і процесів.



Статистична таблиця – це форма раціонального, наочного та систематизованого викладання числових характеристик досліджуваних суспільних явищ і процесів.

У статистичній таблиці за аналогією з граматичним реченням є підмет і присудок.



Підметом статистичної таблиці являється статистична сукупність, тобто ті об'єкти або їх частини, які характеризуються рядом числових показників.

Присудком статистичної таблиці називаються показники, що характеризують досліджувану сукупність і її частини.

Підмет таблиці переважно розташовується зліва і складає зміст рядків.

Присудок таблиці, як правило, розміщується зверху і складає зміст граф. Підмет і присудок інколи можуть мінятися місцями.

Кожна таблиця має ряд горизонтальних рядків і вертикальних граф (стовпчиків, колонок), які при перетині утворюють клітини, котрі заповнюються статистичними даними.

Розграфлена сітка (без речень, слів і цифр) складає скелет таблиці. Якщо записати назви підмета і присудка, матимемо макет таблиці. Для отримання повної таблиці потрібно заповнити усі клітини відповідними статистичними даними. Обов'язковою складовою частиною статистичної таблиці є загальний і внутрішній заголовки. Загальний заголовок розташовується над таблицею, в ньому коротко зазначається, про що йдеться мова в таблиці, до якого місця і часу вона відноситься, в яких одиницях наведені дані.

Внутрішні заголовки розміщуються усередині таблиці (збоку і зверху).

Пристаючи до складання статистичної таблиці, потрібно, перш за все, розробити її макет.

МАКЕТ СТАТИСТИЧНОЇ ТАБЛИЦІ

Назва таблиці (загальний заголовок)

Присудок Предмет	Назва граф (верхні заголовки)								
	1	2	3	4	5	6	7	8	9
А									
Назва рядків (бокові заголовки)									
Підсумок									

Нумерація граф

Рядки таблиці

Підсумковий рядок

Графи таблиці

До деяких таблиць подаються примітки, де роз'яснюється зміст окремих заголовків чи показників.

Залежно від **побудови підмета** розрізняють три види статистичних таблиць: прості, групові, комбінаційні. У простих таблицях в підметі перераховують лише одиниці сукупності, у групових – цифрові дані об'єднують в групи, а у комбінаційних – групи ще розбивають на підгрупи.

■ **Простими** називаються статистичні таблиці, в підметі яких немає груп.

■ Їх ще називають **інформаційними**.

Уся різноманітність простих статистичних таблиць може бути зведена до перелікових, територіальних і хронологічних або до деяких їх сполучень.

- Простими **переліковими** таблицями називаються такі, у підметі яких наводиться перелік одиниць або показників, що вивчаються.
- Простими **територіальними** називаються таблиці, в підметі яких дається перелік територій (країн, областей, районів), кожна з яких характеризується відповідними показниками.
- Простими **хронологічними** називаються таблиці, в підметі яких наводяться певні відрізки часу, а в присудку – один або декілька статистичних показників.
- **Груповими** називаються такі статистичні таблиці, підмет яких утворений в результаті групування одиниць досліджуваного об'єкта за тією чи іншою ознакою.
- **Комбінаційними** називаються статистичні таблиці, у підметі яких групи утворені за однією ознакою діляться на підгрупи за іншими ознаками.

Крім розглянутих вище таблиць, статистика використовує також типові і балансові таблиці.

- **Типовими** називаються статистичні таблиці, підмет яких утворений в результаті типологічного групування.
- **Балансовими** називаються таблиці, які характеризують зв'язок між поступленням і витрачанням ресурсів.

При складанні статистичних таблиць потрібно дотримуватися наступних правил:

1. Кожна статистична таблиця повинна включати тільки ті відмінності, які необхідні для розуміння суті і вивчення даного явища.
2. Заголовки повинні бути сформульовані точно, коротко і ясно.

3. Рядки в підметі і графі в присудку, як правило, повинні нумеруватися.
4. Суворо дотримуватись таких умовних позначень: якщо явище відсутнє, ставиться тире (—); якщо відсутні відомості про його розмір, ставиться три крапки (...) або пишеться „нема відомостей”.
5. Абсолютні дані у межах однієї графі повинні бути округлені з однаковим ступенем точності (до 0,1; 0,01 або 0,001).
6. У випадку великої різноманітності одиниць виміру виділяється спеціальна графа „одиниці виміру”.
7. Таблиці, як правило, повинні бути замкнені, тобто мати підсумкові рядки.

2. Поняття статистичного графіка. Роль графічного методу у статистиці

Статистика за визначенням – це суспільна наука, що властивими тільки їй методами вивчає кількісну сторону масових суспільних явищ у нерозривному зв'язку з їх якісною стороною. Вона дає числові вираження закономірностям розвитку суспільних явищ у конкретних умовах місця й часу. Завданням статистики є збір, аналіз і подання даних.

Ціль графічного методу – отже й статистичних графіків – наочно представити закономірності, виявлені у процесі аналізу статистичної інформації. Графіки у доступній формі, наочно ілюструють виявлені закономірності. Вони значно полегшують сприйняття інформації й пошук закономірностей у даних. Не випадково кажуть, що один графік краще тисячі слів.

Графіки в статистиці мають не тільки ілюстративне значення, вони дозволяють одержати додаткові знання про предмет дослідження, які у цифровому варіанті залишаються схованими, невиявленими. Графічний метод дозволяє виявляти й причинно-наслідкові зв'язки досліджуваних процесів. Результати статистичного дослідження на основі будь-якого

методу (групування, методу середніх, індексного методу, кореляційно-регресійного аналізу тощо) в остаточному підсумку доповнюється використанням графічного методу.

Саме тому усі стандартні статистичні пакети програм для ПК містять потужні графічні блоки, що дозволяють одержати сотні різномунітних графіків, діаграм, включаючи й тривимірні.

Початком використання графічного методу у статистиці вважати роботи шотландського інженера й політеконома У. Плейфейра (William Playfair (1759 – 1823) „Комерційний і політичний атлас” (1786 р.) і „Статистичні огляди” (1801 р.). У цих роботах вперше були використані графічні зображення статистичних даних у вигляді лінійних, секторних і стовпчикових діаграм. Цікаво зазначити, що ім'я головного розробника методів графічного подання статистичних даних перекладається як „той, що грає чесно”.

Статистичний графік – це технічний засіб для наочного зображення статистичних величин за допомогою геометричних образів (геометричні лінії, фігури, сукупність крапок, поверхонь і тощо).

Статистичні графіки – це, головним чином, площинні зображення, на яких статистичні сукупності, що характеризуються певними показниками, описуються за допомогою умовних геометричних образів або знаків.

Геометричні образи дозволяють зіставляти розміри сукупностей, створювати моделі структур, моделі динаміки, розміщення й зв'язки явищ.

Графічний метод є природним продовженням табличного методу. Кожна таблиця може бути представлена графічно.

Статистичні графіки необхідно відрізняти від графіків, які є допоміжними, від графіків математичних.

Подання даних таблиці у вигляді графіка дозволяє втілити сильніше враження, ніж цифри, дозволяє краще осмислити результати статистичного спостереження, правильно їх витлумачити, значно полегшує розуміння статистичного матеріалу, робить його наочним і доступним. Вони дають

нове знання про предмет дослідження, будучи методом узагальнення вихідної інформації.

Статистичні графіки будуються таким чином, щоб при мінімальних зусиллях вони найбільшою мірою полегшували чесний аналіз даних. „Якщо довго катувати дані, вони скажуть вам те, що ви хочете від них почути”.

3. Складові елементи статистичного графіка й правила його побудови

При побудові графічного зображення слід дотримуватися ряду вимог. Насамперед графік повинен бути досить наочним, тому що весь зміст графічного зображення в тому й полягає, щоб наочно відобразити статистичні показники. Крім того, графік повинен бути виразним, дохідливим і зрозумілим. Для виконання перерахованих вище вимог кожний графік повинен включати ряд основних елементів: геометричний образ; поле графіка; просторові орієнтири; масштабні орієнтири; експлікацію графіка.

Якщо розглядати статистичний графік як площинне зображення (двовимірний геометричний знак), то в ньому можна виділити наступні елементи:

Геометричний образ – це символи понять, що відображаються на графіку. Геометричні фігури, сукупність крапок, ліній, фігур, силуетів за допомогою яких зображуються статистичні показники.

Поле графіка – це частина площини, де розташовані геометричні образи. Поле графіка має певні розміри і пропорції сторін.

Просторові орієнтири графіку задаються у вигляді системи координатних сіток. Система координат необхідна для розміщення геометричних знаків у полі графіку.

Масштабні орієнтири статистичного графіка визначаються масштабом і системою масштабних шкал.

Масштаб статистичного графіка – це міра перекладу числової величини в графічну.

Масштабною шкалою називається лінія, окремі крапки якої можуть бути прочитані як певні числа. Шкала має велике значення в графіку й включає три елементи: лінію (або носій шкали), певне число позначених рисками крапок, які розташовані на носії шкали в певному порядку, цифрове позначення чисел, що відповідають окремим позначеним крапкам.

Експлікація графіка – словесне пояснення змісту графіка й значення кожного його геометричного знака.

Правила побудови графіків:

- правильний вибір типу графіка, способу зображення. Графік повинен відповідати змісту. Ніякий графік сам по собі не замінює статистичних даних, тому усі числа повинні бути нанесені на графік;
- графік повинен бути точним і відповідати обраному масштабу, забезпечувати читання графіка та пояснення змісту усіх атрибутів графіка;
- графік не повинен бути перевантажений лініями: не більше 3 – 4 ліній;
- кожен графік повинен мати заголовок і номер. У заголовку необхідно відбити об'єкт, показники, місце й час. Якщо графік розташований у тексті, то заголовок пишуть над графіком; якщо на планшетах – то теж над графіком.

4. Цілі й завдання, що вирішуються за допомогою графіків

Статистичні графіки створюють для:

- 1) порівняння величин між собою (порівняння чисельності населення окремих країн, підприємств по обсягам продукції тощо);
- 2) характеристики складу явища, його структури й структурних зрушень (статевий склад населення, структура основних виробничих фондів, собівартість продукції тощо);
- 3) характеристики розвитку явища в часі (темпи росту

- продуктивності праці тощо);
- 4) вивчення залежності й зв'язку соціально-економічних явищ (залежність випуску продукції від величини основних фондів підприємства тощо);
 - 5) розподілу сукупності за призначенням будь-якої ознаки (розподіл робітників за рівнем заробітної плати тощо);
 - 6) характеристики розвитку явищ у просторі, для економіко-географічної характеристики явищ.

5. Форми й види графіків

Існує безліч видів графічних зображень. Їх класифікація заснована на ряді ознак:

- а) спосіб побудови графічного образу;
- б) геометричні знаки, що зображують статистичні показники;
- в) завдання, розв'язувані за допомогою графічного зображення.

За способом побудови статистичні графіки поділяють на *діаграми* й *статистичні карти*.

Діаграми – найпоширеніший спосіб графічних зображень. Це графіки кількісних відносин. Види й способи їх побудови різноманітні. Діаграми застосовуються для наочного зіставлення в різних аспектах (просторовому, часовому й тощо) незалежних одна від одної величин: територій, населення тощо. При цьому порівняння сукупностей здійснюється за будь-якою істотною ознакою, що варіює.

Залежно від застосовуваних графічних образів серед діаграм розрізняють:

- стовпчикові діаграми;
- площинні діаграми;
- секторні діаграми;
- об'ємні діаграми;
- лінійні графіки тощо.

Діаграми порівняння – співвідношення статистичних сукупностей по ознаці, що змінюється в просторі (стовпчикові й смугові).

Структурні діаграми – діаграми питомих ваг до загального обсягу (стовпчикові й секторні).

Динамічні діаграми – характеризують зміну явища в часі (стовпчикові й смугові) + образотворчі (фігури) + спіральні з радіусом кола + контрольно-планові графіки.

Графіки зв'язку – залежності кореляційно-регресійного аналізу.

Статистичні карти – графіки кількісного розподілу по поверхні. За своєю основною метою вони близько примикають до діаграм і специфічні лише в тому відношенні, що являють собою умовні зображення статистичних даних на контурній географічній карті, тобто показують просторове розміщення або просторову поширеність статистичних даних. Геометричні знаки – крапки, лінії, площини або геометричні тіла.

Картограми – адміністративний або географічний розподіл сукупності (фонові, ізотопи, крапкові).

Картодіаграми – географічний розподіл відображуваного статистикою явища.

Центрограми – статистико-географічні описи.

Усі статистичні графіки можна розподілити на **презентаційні** й **аналітичні**. Перші призначені для швидкого доведення інформації до не дуже підготовленого споживача. Наступні – для полегшення аналізу, вони не обов'язково повинні бути зрозумілі людині без спеціальної підготовки, зате повинні полегшувати ухвалення рішення про правильну обробку даних. Звідси висновок про те, що презентаційні графіки повинні ставитися до типу, що добре знайомий аудиторії, яка їх буде дивитися, а аналітичний графік повинен бути просто зручним візуальним інструментом аналізу – хто хоче, розбереться. Із цієї причини презентаційні графіки звичайно ставляться до основних типів, описаних ще Плейфейром: стовп-

чикові й стрічкові діаграми, лінійні графіки й кругові діаграми, а також діаграми розсіювання. Кількість аналітичних графіків значно більше – це й коробчасті графіки, і діаграми впливу (див. регресійний аналіз), і чотирипільні діаграми, і мозаїчні графіки, і діаграми-сита і багато інших.

Для подання кожного типу даних повинен використовуватися певний тип графіка:

- Кількість спостережень у декількох групах – стовпчикова діаграма (гістограма);
- Середні значення змінної в декількох групах – стовпчикова діаграма;
- Середні значення змінної при наявності великої кількості груп – смугова діаграма;
- Зміна значення параметру в часі – лінійна діаграма;
- Процентний розподіл в одній групі – кругова діаграма;
- Процентний розподіл у декількох групах – процентна стовпчикова діаграма;
- Перебіг змінної циклічного характеру (дні тижня, пори року, місяці) – зірчаста діаграма;
- Зв'язок двох змінних одна з одним – діаграма (графік) розсіювання.

Багато відомих графіків є похідними цих базових типів графіків.



Питання для самоконтролю по темі 12

1. *Поняття про статистичні таблиці.*
2. *Вимоги до побудови статистичних таблиць.*
3. *Призначення статистичних таблиць.*
4. *Завдання табличного методу викладу статистичних даних.*
5. *Підмет та присудок статистичних таблиць.*
6. *Поняття про статистичні графіки.*
7. *Вимоги до побудови статистичних графіків.*
8. *Призначення статистичних графіків та графічного методу.*
9. *Основні елементи статистичних графіків.*
10. *Що таке шкала статистичного графіку?*
11. *Загальні поняття про види статистичних графіків.*
12. *Використання діаграм.*
13. *Контрольно-планові графіки.*
14. *Картограми і картодіаграми.*

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

Основна навчальна література

1. Адамов В.Е. Факторный индексный анализ. – М.: Статистика, 1987. – 199 с.
2. Герчук Я.П. Графические методы в статистике. – М.: Статистика, 1972. – 78 с.
3. Елисеева И.И., Юзбашев М.М. Общая теория статистики: Учебник – М.: Финансы и статистика, 1996. – 368 с.
4. Ефимова М.Р., Петрова Е.В., Румянцев В.Н. Общая теория статистики: Учебник. – М.: ИНФРА. – М. 1998. – 416 с.
5. Кендэл М. Ранговые корреляции. – М.: Статистика, 1975. – 241 с.
6. Кильдишев Г.С., Аболонцев Ю.И. Многомерные группировки. – М.: Статистика, 1978. – 160 с.
7. Кильдишев Г.С., Овсиенко В.Е., Рабинович П.М., Рябушкин Т.В. Общая теория статистики: Учебник – М.: Статистика, 1980. – 423 с.
8. Кильдишев Г.С., Френкель А.А. Анализ временных рядов и прогнозирование. – М.: Статистика, 1973. – 103 с.
9. Ковалевский Г.В. Индексный метод в экономике. – М.: Финансы и статистика, 1989. – 239 с.
10. Ковалевский Г.В. Идеи, поиски, решения. Харьковская экономическая школа (1804-2004). Монография. – Харьков: ХНАГХ, 2005. – 179 с.
11. Ковалевский Г.В. Индексный метод в социальной статистике // Проблемы социальной статистики. – М.: Наука, 1986. С. 151-162.
12. Ковалевский Г.В. Основные проблемы современной статистики: новые идеи, поиски, решения // Теорія і методологія статистичного аналізу. – К.: КНЕУ, 2006. С. 392-394.
13. Ковалевский Г.В. Статистика. Підручник. – Харьков: ХНАГХ, 2010. – 313 с.

14. Общая теория статистики / Под ред. А.Я.Боярского и Г.Я.Громыко. 2-е изд. - М.: МГУ, 1985. – 376 с.
15. Общая теория статистики: Учебник / Г.С.Кильдишев, В.Е.Овсиенко, П.М.Рабинович, Т.В.Рябушкин. – М.: Статистика, 1980. – 423 с.
16. Пасхавер И.С. Средние величины в статистике. – М.: Статистика, 1979. – 279 с.
17. Рязузов Н.Н. Общая теория статистики: Учебник. – М.: Финансы и статистика, 1984. – 343 с.
18. Статистика: Підручник / А.В.Головач, А.М.Єріна, О.В.Козирев та ін. / Під ред. А.В.Головача. – К.: Вища школа, 1993. – 623 с.: іл.
19. Статистика: Підручник / С.С.Герасименко, А.В.Головач, А.М.Єріна та ін. / Під наук. ред. д-ра екон. наук С.С.Герасименка. – 2-е вид., перероб. і доп. – К.: КНЕОІ, 2000. – 467 с.
20. Статистический словарь. 3-е изд. – М.: Финансы и статистика, 1996.- 642 с.
21. Теория статистики: Учебник/ под ред. Р.А. Шмойловой. 4-е изд. – М.: Финансы и статистика, 2004. – 467 с.
22. Теорія статистики: Навчальний посібник / Вашків П.Г., Пастер П.І., Сторожук В.П., Ткач Є.І. – К.: Либідь, 2001. – 320 с.
23. Ткач Є.І. Загальна теорія статистики: Підручник. – Тернопіль: Лідер, 2004. – 388 с.
24. Уманець Т.В., Пігарєв Ю.Б. Статистика: Навч. посіб. – К.: Вікар, 2003. – 623с.
25. Урланис Б.Ц. Общая теория статистики: Учебник. – М.: Статистика, 1973. – 439 с.
26. Четыркин Е.М. Статистические методы прогнозирования. – М.: Статистика, 1977. – 200 с.

Методичне забезпечення

27. Сніжко С.В. Програма та робоча програма навчальної дисципліни «Статистика» (для студентів 2 курсу денної та 3 курсу заочної форми навчання освітньо-кваліфікаційного рівня бакалавр, напряму підготовки 0501 – «Економіка і підприємництво», спеціальності 6.050100 «Облік і аудит»). – Х.: ХНАМГ, 2009. – 28с.

Internet ресурси

28. Державний комітет статистики // www.ukrstat.gov.ua
29. Головне управління статистики в Харківській області // www.uprstat.kharkov.ukrtel.net
30. Цифровий репозиторій ХНАМГ // eprints.ksame.kharkov.ua

НАВЧАЛЬНЕ ВИДАННЯ

СНІЖКО Сергій Володимирович

Конспект лекцій з курсу «**Статистика**»(для студентів 2 курсу денної та 3 курсу заочної форми навчання освітньо-кваліфікаційного рівня бакалавр, галузі знань 0305 «Економіка і підприємництво», напряму підготовки 6.030509 «Облік і аудит»).

Редактор *М. З. Аляб'єв*
Комп'ютерне верстання *С. В. Сніжко*

План 2010, поз. 140Л

Підп. до друку 10.09.2010
Друк на ризографі.
Зам. №

Формат 60x84 1/16
Ум. друк. арк. 8,2
Тираж 50 пр.

Видавець і виготовлювач:
Харківська національна академія міського господарства,
вул. Революції, 12, Харків, 61002
Електронна адреса: rectorat@ksame.kharkov.ua
Свідоцтво суб'єкта видавничої справи: ДК №731 від 19.12.2001